

# Correlated Fluctuations in Strongly Coupled Binary Networks Beyond Equilibrium

David Dahmen,<sup>1</sup> Hannah Bos,<sup>1</sup> and Moritz Helias<sup>1,2</sup><sup>1</sup>*Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA BRAIN Institute I, Jülich Research Centre, 52425 Jülich, Germany*<sup>2</sup>*Department of Physics, Faculty I, RWTH Aachen University, 52074 Aachen, Germany*

(Received 9 December 2015; revised manuscript received 24 May 2016; published 19 August 2016)

Randomly coupled Ising spins constitute the classical model of collective phenomena in disordered systems, with applications covering glassy magnetism and frustration, combinatorial optimization, protein folding, stock market dynamics, and social dynamics. The phase diagram of these systems is obtained in the thermodynamic limit by averaging over the quenched randomness of the couplings. However, many applications require the statistics of activity for a single realization of the possibly asymmetric couplings in finite-sized networks. Examples include reconstruction of couplings from the observed dynamics, representation of probability distributions for sampling-based inference, and learning in the central nervous system based on the dynamic and correlation-dependent modification of synaptic connections. The systematic cumulant expansion for kinetic binary (Ising) threshold units with strong, random, and asymmetric couplings presented here goes beyond mean-field theory and is applicable outside thermodynamic equilibrium; a system of approximate nonlinear equations predicts average activities and pairwise covariances in quantitative agreement with full simulations down to hundreds of units. The linearized theory yields an expansion of the correlation and response functions in collective eigenmodes, leads to an efficient algorithm solving the inverse problem, and shows that correlations are invariant under scaling of the interaction strengths.

DOI: [10.1103/PhysRevX.6.031024](https://doi.org/10.1103/PhysRevX.6.031024)Subject Areas: Biological Physics,  
Complex Systems, Statistical Physics

## I. INTRODUCTION

Understanding collective phenomena arising from the interactions in a many-body system is the challenging subject of many-particle physics. The characterization of the emerging states typically rests on the quantification of correlations [1]. Among the simplest classical models is the Ising (binary) model [2], or Glauber dynamics [3] in its kinetic formulation. While for symmetric random couplings [4] these systems are within the realm of equilibrium statistical mechanics [5], the asymmetric kinetic Ising model [6–8], even in the stationary state, does not reach thermodynamic equilibrium. In their Markovian formulation [9], these processes are studied in the field of nonequilibrium stochastic thermodynamics (for review, see Chap. 6 of Ref. [10]). The methods derived in these fields have proven useful in a variety of disciplines, including computer science, biology, artificial intelligence, social sciences, and economics (see, e.g., Refs. [11–16] and references therein).

Neuronal networks of the central nervous system are prominent examples of nonequilibrium systems due to Dale's principle [17,18], which states that a neuron either

excites or inhibits all its targets. The correlations between the activities of nerve cells are functionally important [19]: They influence the representation of information in population signals depending on the readout in either a detrimental [20] or beneficial [21] manner, their time-dependent modulations are linked to behavior [22], and they determine the influence of neuronal activity on synaptic plasticity [23–25], the biophysical mechanism underlying learning. In the case of binary units (Ising spins), knowledge of pairwise covariances, moreover, proves useful for sampling-based inference (reviewed in Ref. [26]) as they constitute the next order in the systematic expansion of the joint probability distribution beyond independence [cf. Eq. (22) in Ref. [3]].

Dynamic mean-field theory [27–32] in the  $N \rightarrow \infty$  limit effectively reduces the many-body problem of a network comprised of a large number of units to a single particle interacting with a self-consistently determined field, but it neglects the cross-covariances of activities between units. Ginzburg and Sompolinsky [33] introduced pairwise correlations to the description of weakly coupled systems, and Renart *et al.* [34] extended the analysis to strongly coupled units in the large- $N$  limit. Both approaches are, however, limited to averaged pairwise correlations. Similarly, approximate master equations for binary-state dynamics on complex networks, as recently discussed in context of the pair approximation

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

by Gleeson [35,36], are restricted to global dynamics in infinite networks. For self-averaging observables, Coolen *et al.* [37] presented a systematic approach to derive closed equations of motion, which, however, are only valid in the limit  $N \rightarrow \infty$ .

In the current work, we derive a systematic cumulant expansion yielding an analytical description of correlations between the activities of individual pairs of units, which goes beyond averaged pairwise correlations [33,34,38] and is applicable to a single realization of an asymmetric (directed) network. The framework links the coupling structure to the emerging correlated activity and describes the first- and second-order statistics for single units and pairs of units in a large, but finite, network of possibly strongly interacting elements. The analytical expressions predict a distribution of pairwise correlations with a small mean but large standard deviation, as observed experimentally [39] in neural tissue and yet not explained theoretically. Moreover, the framework can be employed to infer the couplings between the units from the observed correlated activity, also termed the inverse problem [40–43].

In particular, we show that a Gaussian truncation of the presented cumulant hierarchy already yields good predictions for individual mean activities and pairwise correlations. Taking into account the subset of third-order cumulants that for binary variables can be expressed by lower-order ones improves the prediction significantly. This finding demonstrates that the second-order statistics suffices to capture the major features of the collective dynamics arising in random binary networks, even when the coupling is strong. Our approach consistently takes into account the network-generated fluctuations in the marginal statistics of the input to each unit in a similar spirit as the seminal work by van Vreeswijk and Sompolinsky [44]. This approach exposes peculiar features of networks of binary threshold units. For a fixed number of incoming connections to each unit, mean-field theory predicts their averaged activities to be identical. Here, we show how distributed mean activities arise solely from the correlations emerging in the network. Units with a hard activation threshold render covariances independent of synaptic amplitudes  $J$ : If all incoming connections as well as the threshold of a unit are changed proportionally, the mean activity and correlations are maintained. Hence, scaling the threshold appropriately, it is impossible to increase the influence of one unit on another by larger coupling amplitudes. This finding questions the customary division into strong ( $J \propto N^{-(1/2)}$ ) and weak coupling ( $J \propto N^{-1}$ ) in such systems. Here, we show, in addition, that a network with hard threshold units implements the strongest possible coupling between stochastic binary units.

The independence of covariances with respect to coupling strengths implies that the latter cannot be reconstructed from the knowledge of the activity alone.

However, together the amplitude and slope of covariance functions at zero time lag uniquely determine a linearized effective coupling strength between units. A linear approximation of the dynamics of fluctuations around the stationary state leads to a modified Lyapunov equation. Decomposing the fluctuations into characteristic eigenmodes of the network shows that, to linear order, the kinetic binary network is equivalent to a system of coupled Ornstein-Uhlenbeck processes [45,46]. The linear description further allows for the decomposition of the response to external stimulations into eigenmodes of the system.

## II. FIRST AND SECOND MOMENTS OF THE JOINT PROBABILITY DISTRIBUTION

Here, we consider the classical network model of stochastic binary units and denote the activity of unit  $k$  as  $n_k(t)$  being either 0 or 1, where 1 indicates the active and 0 the inactive state (see, e.g., Refs. [33–36,38,44,47–49]). Since we may interpret a unit as representing an individual neuron, we use the terms “unit” and “neuron” interchangeably. The model neurons show transitions at random points in time between the two states 0 and 1 controlled by transition probabilities. The state of the network of  $N$  such units is described by a binary vector  $\mathbf{n} = (n_1, \dots, n_N) \in \{0, 1\}^N$  that at each time point assumes one of the  $2^N$  possible states. Using an asynchronous update [50], in each infinitesimal interval  $[t, t + \delta t)$ , each unit in the network has the probability  $(1/\tau)\delta t$  to be chosen for an update [51], where  $\tau$  is the time constant of the dynamics. An equivalent implementation draws the time points of an update independently for all units. For a particular unit, the sequence of update points has exponentially distributed intervals with mean duration  $\tau$ , i.e., the update times form a Poisson process with rate  $\tau^{-1}$ . We employ the latter implementation in the globally time-driven [52] spiking simulator NEST [53] as described in Grytskyy *et al.* (see Appendix of Ref. [54]) and use a discrete time resolution  $\delta t = 0.1$  ms for the intervals. The stochastic update constitutes a source of noise in the system. Given that the  $k$ th neuron is selected for update, the probability to end in the active state ( $n_k = 1$ ) is determined by the activation function  $F_k(\mathbf{n})$ , which depends on the activity  $\mathbf{n}$  of all units  $i$  that are connected to neuron  $k$  via couplings  $J_{ki}$ . The probability to end in the inactive state ( $n_k = 0$ ) is given by  $1 - F_k(\mathbf{n})$ .

The stochastic system at time  $t$  is completely determined by its probability distribution  $p(\mathbf{n}, t)$ . The time evolution of the two-point distribution  $p(\mathbf{n}, t, \mathbf{q}, s)$  (describing the probability that the system was in state  $\mathbf{q}$  at time  $s$  and is in state  $\mathbf{n}$  at time  $t$ ) obeys—because of the Markov property—the same master equation [see Eq. (A1) in Appendix A] as  $p(\mathbf{n}, t)$ . Using the definitions of the first moment and the equal time, as well as the two time point second moments,

$$\begin{aligned}
\langle n_k(t) \rangle &:= \sum_{\mathbf{n}} p(\mathbf{n}, t) n_k, \\
\langle n_k(t) n_l(t) \rangle &:= \sum_{\mathbf{n}} p(\mathbf{n}, t) n_k n_l, \\
\langle n_k(t) n_l(s) \rangle &:= \sum_{\mathbf{n}, \mathbf{q}} p(\mathbf{n}, t, \mathbf{q}, s) n_k q_l,
\end{aligned} \tag{1}$$

one obtains a set of differential equations for the first and second moments [33,34,38,55], which read [for completeness, in the Appendix, we included their derivation in Eq. (A4)]

$$\begin{aligned}
\tau \frac{\partial}{\partial t} \langle n_k(t) \rangle &= -\langle n_k(t) \rangle + \langle F_k(\mathbf{n}(t)) \rangle, \\
k \neq l \quad \tau \frac{\partial}{\partial t} \langle n_k(t) n_l(t) \rangle &= -2\langle n_k(t) n_l(t) \rangle \\
&\quad + \langle F_k(\mathbf{n}(t)) n_l(t) \rangle \\
&\quad + \langle F_l(\mathbf{n}(t)) n_k(t) \rangle, \\
t > s \quad \tau \frac{\partial}{\partial t} \langle n_k(t) n_l(s) \rangle &= -\langle n_k(t) n_l(s) \rangle \\
&\quad + \langle F_k(\mathbf{n}(t)) n_l(s) \rangle,
\end{aligned} \tag{2}$$

where the expectation value  $\langle \rangle$  defined in Eq. (1) can be interpreted as an average over realizations of the random dynamics. Note that the second line does not hold for  $k = l$  but becomes the first line because of  $\langle n_k^2 \rangle = \langle n_k \rangle =: m_k$ . The third line becomes the second line for  $t \rightarrow s$ . These equations are identical to Eqs. (3.4)–(3.7) in Ref. [33], to Eqs. (3.12) and (3.13) in Ref. [38], and to Eqs. (18)–(22) in the Supplementary Material of Ref. [34]. The works by Ginzburg and Sompolsky [33] and Renart *et al.* [34] considered the second-order statistics averaged over many pairs of neurons, i.e.,  $\langle n_\alpha n_\beta \rangle = 1/(N_\alpha N_\beta) \sum_{i \in \alpha, j \in \beta} \langle n_i n_j \rangle$ . These averages are closely related to the population-averaged activities  $n_\alpha(t) = (1/N_\alpha) \sum_{i \in \alpha} n_i(t)$  and can therefore be treated by population-averaging mean-field methods. While the methodology in Buice *et al.* [38] is general, the authors consider interacting populations of neurons. Here, we go beyond the population-level description and consider second-order statistics of individual pairs for a particular coupling matrix  $\mathbf{J}$ , a setting which is beyond the disorder-averaged description of self-averaging observables [37]. In particular, we are interested in the stationary statistics of the mean activities of individual units and their zero time lag pairwise covariances  $c_{kl} = \langle n_k(t) n_l(t) \rangle - \langle n_k(t) \rangle \langle n_l(t) \rangle$ , which can be deduced from Eq. (2),

$$\begin{aligned}
m_k &= \langle F_k(\mathbf{n}) \rangle, \\
c_{kl} &= \begin{cases} \frac{1}{2} \langle F_k(\mathbf{n}) n_l \rangle + \frac{1}{2} \langle F_l(\mathbf{n}) n_k \rangle - m_k m_l & k \neq l \\ m_k (1 - m_k) & k = l. \end{cases}
\end{aligned} \tag{3}$$

We now assume a standard form for the activation function between neurons given by

$$\begin{aligned}
F_k(\mathbf{n}) &= f(h_k(\mathbf{n})), \\
h_k &= \sum_{i=1}^N J_{ki} n_i,
\end{aligned} \tag{4}$$

where the gain function  $F_k(\mathbf{n}) = f(h_k(\mathbf{n}))$  depends on the state of the network only via the summed and weighted scalar activity  $h_k$ . In systems with symmetric couplings  $J_{ki} = J_{ik}$  and activation function  $f = \tanh$ , this choice describes a system with a Hamiltonian, which is quadratic in the state variables and whose stationary distribution is of Boltzmann form. Instead, here we study general (nonsymmetric) couplings and arbitrary gain functions  $f$ . However, for concreteness, when comparing the analytical predictions to numerical results, we choose  $f(h_k) = H(h_k - \theta)$ , where  $H(x) = 1$  for  $x \geq 0$  and 0 else is the Heaviside function. Intermediate results also hold for arbitrary gain functions  $f$ . If the distribution of  $h_k$  was known, the expectation value  $\langle F_k \rangle$  could be directly calculated.

### III. GAUSSIAN APPROXIMATION

We aim at a self-consistent equation for the  $N$  values of the first moments and the  $N(N-1)$  values of the second moments of the activity variables (3). We note that van Vreeswijk *et al.* [44] and Renart *et al.* [34] invoke the central-limit theorem to justify the assumption that the local field  $h_k$  typically follows a Gaussian distribution  $\mathcal{N}(\mu_k, \sigma_k^2)$  with cumulants  $\mu_k$  and  $\sigma_k^2$ . These cumulants are related to cumulants of the activity variables  $\mathbf{n}$  via

$$\begin{aligned}
\mu_k &= \langle h_k \rangle = \sum_i J_{ki} \langle n_i \rangle = (\mathbf{Jm})_k, \\
\sigma_k^2 &= \langle h_k^2 \rangle - \mu_k^2 = \sum_{i,j} J_{ki} J_{kj} (\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle) \\
&= (\mathbf{JCJ}^T)_{kk},
\end{aligned} \tag{5}$$

with the vector of mean activities  $m_i = \langle n_i \rangle$  and the covariance matrix  $c_{ij} = \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle$ , which contains  $c_{ii} = \langle n_i \rangle (1 - \langle n_i \rangle)$  on the diagonal. In the seminal work by van Vreeswijk and Sompolsky [56], the influence of the cross-covariances on the variance in the input to a neuron was neglected. Instead, only the dominant contribution, given by the variances  $c_{ii}$  of the single neurons on the diagonal, was taken into account, leading to the expression  $\sigma_k^2 = \sum_i J_{ki}^2 m_i (1 - m_i)$ . The additional dependence of  $\sigma_k^2$  on the off-diagonal elements (5) is important for the distribution of mean activities, as we show in the following. We note that because of the marginal binary statistics of  $n_i$ , it follows that  $\langle n_i^2 \rangle = \langle n_i \rangle$ , showing that the second moment is uniquely determined by the first moment. We exploit this property in the subsequent sections by expressing a subset of higher-order cumulants in terms of lower-order ones. Using the central-limit theorem, as outlined above, leads to

$$m_k = \langle F_k(\mathbf{n}) \rangle \simeq \int_{-\infty}^{\infty} H(x - \theta) \mathcal{N}(\mu_k, \sigma_k^2, x) dx \\ = \frac{1}{2} \operatorname{erfc} \left( -\frac{\mu_k - \theta}{\sqrt{2}\sigma_k} \right). \quad (6)$$

To enable the extension to further corrections, we now aim at a more systematic calculation of expectation values containing the gain function. Using the Fourier representation  $f(x) = (1/2\pi) \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \exp(i\omega x)$  of the gain function  $f$ , we obtain

$$\langle F_k(\mathbf{n}) \rangle_{\mathbf{n}} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \left\langle \exp \left( (i\omega) \sum_{j=1}^N J_{kj} n_j \right) \right\rangle_{\mathbf{n}} \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \langle \exp((i\omega) h_k(\mathbf{n})) \rangle_{\mathbf{n}}, \quad (7)$$

where we inserted the definition of  $h_k$  [Eq. (4)] in the second line. Defining the vector  $\mathbf{t}$  with elements  $t_j = i\omega J_{kj}$ , the expectation value in the first line of the previous expression has the form  $\langle \exp(\mathbf{t} \cdot \mathbf{n}) \rangle_{\mathbf{n}}$ , which is the definition of the characteristic function or moment generating function

$$\varphi_{\mathbf{n}}(\mathbf{t}) \equiv \langle e^{\mathbf{t} \cdot \mathbf{n}} \rangle_{\mathbf{n}} \\ = \sum_{\mathbf{n} \in \{0,1\}^N} p(\mathbf{n}) e^{\mathbf{t} \cdot \mathbf{n}} \\ = \exp(\Phi_{\mathbf{n}}(\mathbf{t})) \quad (8)$$

of the joint distribution of the binary variables  $p(\mathbf{n})$  (see page 32 of Ref. [57]), which can also be expressed by the cumulant generating function  $\Phi_{\mathbf{n}}(\mathbf{t})$ . Note that for compactness of notation, we omitted the index  $k$  for  $\mathbf{t}$ . The Taylor expansion of  $\Phi_{\mathbf{n}}(\mathbf{t})$  explicitly introduces a cumulant hierarchy, showing that cumulants of activity variables  $n_j$  at all orders contribute to  $\langle F_k(\mathbf{n}) \rangle$ . If we neglect cumulants higher than order two, thus effectively approximating the binary states as correlated Gaussian variables, we use the corresponding quadratic cumulant generating function

$$\Phi_{\mathbf{n}}(\mathbf{t}) = \sum_i m_i t_i + \frac{1}{2} \sum_{i,j} c_{ij} t_i t_j \quad (9)$$

and obtain

$$\langle F_k(\mathbf{n}) \rangle_{\mathbf{n}} \simeq \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \exp \left( \sum_i m_i t_i + \frac{1}{2} \sum_{i,j} c_{ij} t_i t_j \right) \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \exp \left( \mu_k i\omega + \frac{1}{2} \sigma_k^2 (i\omega)^2 \right). \quad (10)$$

Here, we identified the sums in the penultimate line as the mean and variance of the input field (5). On the other hand, in the second line of Eq. (7), we identify the moment

generating function  $\varphi_{h_k}(i\omega) \equiv \langle \exp((i\omega) h_k(\mathbf{n})) \rangle_{\mathbf{n}}$  of the input field  $h_k$ . From Eq. (10), we obtain its corresponding cumulant generating function as  $\Phi_{h_k}(i\omega) = \mu_k i\omega + \frac{1}{2} \sigma_k^2 (i\omega)^2$ . We therefore conclude that  $h_k \sim \mathcal{N}(\mu_k, \sigma_k^2)$  follows a Gaussian distribution. The mean activity is then given by

$$m_k = \langle F_k(\mathbf{n}) \rangle_{\mathbf{n}} = \langle f(h_k) \rangle_{h_k \sim \mathcal{N}(\mu_k, \sigma_k^2)}, \quad (11)$$

which for  $f(x) = H(x - \theta)$  is identical to Eq. (6). The procedure hence reproduces the known result for the Heaviside gain function (see Supplementary Material, Secs. 2.2 and 2.3 of Ref. [34,58]) but additionally yields a generalization for smooth gain functions  $f$  that takes into account the fluctuations of the synaptic input, which is in line with the treatment in Eq. (27) of Ref. [59].

More importantly, this systematic calculation reveals the assumption that is underlying the Gaussian approximation for the input field, namely, that cumulants higher than order two are ignored on the level of individual neuronal activities.

To obtain an expression for the zero time lag covariance, we start from Eq. (3). We hence need to evaluate expressions of the form  $\langle F_k(\mathbf{n}) n_l \rangle$  which do not factorize trivially since the value of  $n_l$  may have an influence on neuron  $k$  through the gain function  $F_k(\mathbf{n})$ . However, along the same lines as above, we can derive  $\langle F_k(\mathbf{n}) n_l \rangle$  using the identical approximation of  $\langle \exp(\mathbf{n} \cdot \mathbf{t}) \rangle$  by first noting

$$\langle F_k(\mathbf{n}) n_l \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \left\langle \exp \left( i\omega \sum_{j=1}^N J_{kj} n_j \right) n_l \right\rangle \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \hat{f}(\omega) \frac{\partial}{\partial t_l} \langle \exp(\mathbf{n} \cdot \mathbf{t}) \rangle.$$

Again expressing the characteristic function by the cumulant generating function (9) yields

$$\frac{\partial}{\partial t_l} \exp(\Phi_{\mathbf{n}}(\mathbf{t})) = \underbrace{\frac{\partial \Phi_{\mathbf{n}}(\mathbf{t})}{\partial t_l}}_{=m_l + \sum_j J_{kj} c_{jl}(i\omega) \text{ by (9)}} \exp(\Phi_{\mathbf{n}}(\mathbf{t})) \\ = \left( m_l + \sum_j J_{kj} c_{jl} \frac{\partial}{\partial \mu_k} \right) \exp(\Phi_{h_k}(i\omega)), \quad (12)$$

where we generated the factor  $i\omega$  by a derivative  $\partial_{\mu_k}$  in the last line. With Eq. (11), this yields

$$\langle F_k(\mathbf{n}) n_l \rangle = m_l \langle f(h_k) \rangle_{h_k \sim \mathcal{N}(\mu_k, \sigma_k^2)} + S_k(\mathbf{JC})_{kl}, \\ S_k = \frac{\partial}{\partial \mu_k} \langle f(h_k) \rangle_{h_k \sim \mathcal{N}(\mu_k, \sigma_k^2)} \\ \stackrel{\text{i.b.p.}}{=} \langle f'(h_k) \rangle_{h_k \sim \mathcal{N}(\mu_k, \sigma_k^2)}, \quad (13)$$



where we defined the local susceptibility  $S_k$  that determines the influence of an input to neuron  $k$  onto its output at the time point of its update. For a differentiable gain function  $f$ , the susceptibility is equal to the slope  $f'$  averaged over the fluctuations of  $h_k$  [compare also Eq. (27) of Ref. [59]], which follows from integrating the second line (13) by parts (i.b.p.). Note that the second line in Eq. (13) also holds for nondifferentiable  $f$ , such as our particular choice  $f(x) = H(x - \theta)$ , for which the susceptibility has the form

$$S_k := \frac{1}{\sqrt{2\pi}\sigma_k} e^{-(\mu_k - \theta)^2 / (2\sigma_k^2)}, \quad (14)$$

and which, by Eq. (13), is the strongest possible coupling for a given input statistics. The self-consistent set of equations for the first and second cumulants thus reads

$$m_k = \frac{1}{2} \operatorname{erfc} \left( -\frac{\mu_k - \theta}{\sqrt{2}\sigma_k} \right), \quad (15)$$

$$c_{kl} = \begin{cases} \frac{1}{2} S_k (\mathbf{J}\mathbf{C})_{kl} + \frac{1}{2} S_l (\mathbf{J}\mathbf{C})_{lk} & k \neq l \\ m_k (1 - m_k) & k = l. \end{cases}$$

To obtain the joint solution of Eqs. (15), we use a damped fixed-point iteration with the  $n$ th iteration value denoted as  $(m_k^{(n)}, c_{kl}^{(n)})$ , which has the form

$$m_k^{(n+1)} = \rho \frac{1}{2} \operatorname{erfc} \left( -\frac{\mu_k^{(n)} - \theta}{\sqrt{2}\sigma_k^{(n)}} \right) + (1 - \rho) m_k^{(n)},$$

$$c_{kl}^{(n+1)} = \rho \frac{1}{2} (S_k^{(n)} (\mathbf{J}\mathbf{C}^{(n)})_{kl} + S_l^{(n)} (\mathbf{J}\mathbf{C}^{(n)})_{lk}) + (1 - \rho) c_{kl}^{(n)},$$

$$c_{kk}^{(n+1)} = m_k^{(n+1)} (1 - m_k^{(n+1)}), \quad (16)$$

where we used the damping factor  $\rho = 0.7$ , which, for the current application, yields good convergence. Here,  $\mu_k$  and  $\sigma_k$  are determined by Eq. (5) in each step. The iteration continues until the summed absolute change of  $m_k$  and  $c_{kl}$  is smaller than the chosen tolerance of  $10^{-14}$ , allowing a sufficient safety margin towards truncation errors induced by the employed double precision numerics. Figure 1 shows that the theoretical prediction yields mean activities and covariances in line with simulations of a random fixed-in-degree network of excitatory and inhibitory neurons. While cross-covariances are explained with good accuracy, mean activities slightly, but systematically, deviate from simulated results: The scatter plot, showing the mean activities of individual neurons, has a slope below unity, indicating that the width of the distribution is underestimated by the theory. Moreover, we observe that the population-averaged covariances are slightly underestimated by the theory. However, in summary, the theory in the Gaussian approximation captures not only the mean

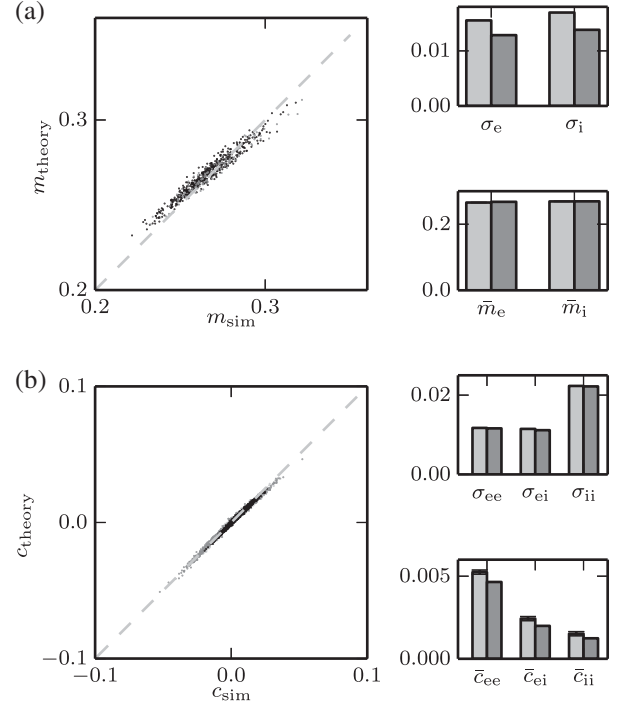


FIG. 1. Consistent Gaussian approximation for the mean activity and covariance matrix in a random network of binary neurons. (a) Theoretical prediction of mean activities (15) versus simulation results (black is for excitatory; gray is inhibitory). The diagonal is indicated by the dashed line. Lower inset: Mean activity averaged over excitatory and inhibitory neurons (light gray is for simulation; darker gray is for theory). Black error bars show standard error of the mean obtained from 20 simulations (not visible, below line width). Upper inset: Width of distribution of mean activities. (b) Theoretical prediction (15) versus simulated covariance. Only cross-covariances are shown, as autocovariances follow trivially by Eq. (3) from the mean activities. Lower inset: Mean covariances averaged over excitatory-excitatory, excitatory-inhibitory, and over inhibitory-inhibitory neuron pairs (light gray is for simulation; darker gray is for theory). Error bars show the standard error of the mean obtained from 20 simulations. Upper inset: Width of the distributions of covariances. Network parameters: Number of excitatory ( $N_E = 500$ ) and inhibitory neurons ( $N_I = 125$ ), threshold  $\theta = -5.5$ , excitatory coupling strength  $J_{ij} = J = 1$ ,  $\forall j \in E$ , inhibitory coupling strength  $J_{ij} = -6J$ ,  $\forall j \in I$ , and fixed in-degree homogeneous random network with connection probability  $p = 0.2$ . Results are averaged over 20 simulations with  $T = 2,000,000$  ms each. Simulations were performed using the NEST simulator [53]. Theoretical predictions are obtained by damped fixed-point iteration (16).

and width of the covariance distribution to a large extent but also its general shape as shown in Fig. 2.

The systematic calculation presented here shows that the Gaussian assumption on the level of the individual statistics directly yields the linearized result of Secs. 2.3 of the Supplementary Material of Ref. [34]. Originally, the terms of the form  $\langle F_k(\mathbf{n}) n_l \rangle$  in Eq. (3) were obtained by using

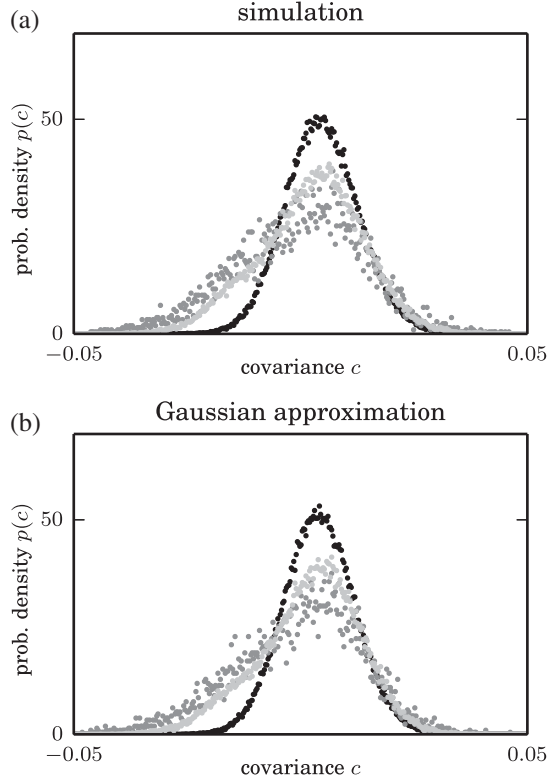


FIG. 2. Distribution of covariances in the Gaussian approximation for a random network of binary neurons. (a) Distribution of cross-covariance between excitatory neurons (black) and inhibitory neurons (gray), and between one inhibitory and one excitatory neuron (light gray) from simulation. (b) Same as (a), but showing the theoretical result for the Gaussian approximation (15). Parameters are the same as in Fig. 1.

relations between conditioned and unconditioned probability distributions of binary variables. After the Gaussian approximation, the resulting nonlinear equation was linearized and second-order terms were discussed to be small in the large- $N$  limit. In contrast, the derivation here shows that the linearization is not an additional assumption but appears naturally once the Gaussian approximation has been performed on the level of individual variables. Note also that no additional assumption regarding the strength of the coupling is necessary.

#### IV. THIRD-ORDER CUMULANTS

The derivation in the previous section is systematic in the sense that all cumulants of order three and higher of the stochastic variables  $\mathbf{n}$  are consistently neglected. The original idea of a mean-field description, however, seeks an approximation of the local field  $h_k$  sensed by an individual unit  $k$  rather than a truncation of the cumulants of the individual variables  $\mathbf{n}$  themselves. Since the local field is, by Eq. (4), a superposition of a large number of weakly correlated contributions, the distribution of  $h_k$  will be close to Gaussian by the central limit theorem. In the

path-integral formulation, commonly employed to study disordered systems, the Gaussian approximation of  $h_k$  is equivalent to a saddle-point approximation to lowest order in the auxiliary field [see, e.g., Eq. (3.5) of Ref. [28] or Eq. (3) of Ref. [29]]. Even though one may expect the central-limit theorem to be applicable to the summed input  $h_k$ , it is easy to see that for networks of several hundreds of units, higher order cumulants still have an effect. This is illustrated in Fig. 3, showing the distribution of the input to a neuron receiving a sum of binary, uncorrelated signals. The Gaussian approximation, with the same mean and variance as the exact distribution, has a peak that is slightly shifted to the left. Taking into account the third-order cumulant cures the displacement of the peak. The expansion in cumulants in Sec. III shows how the higher-order cumulants can be taken into account in the calculation of the expectation values of the gain function.

In the following, we expand the distribution of  $h_k$  up to third-order cumulants. To this end, we use the relationship that the  $n$ th cumulant of a summed variable is the sum of  $n$ th cumulants of its constituents [Eq. (B4)]. We then use the property of binary variables that  $n_k^K = n_k$  for  $K \geq 1$ , which relates cumulants of order  $K + L$  to cumulants of order  $L + 1$  in cases where neuron  $k$  appears  $K$  times in the cumulant (see Appendix D for the detailed calculation). Note that similar relations also hold for classical spins  $n_i = \pm 1$ , or generally for dichotomous units with arbitrary states  $n_i \in \{\alpha, \beta\}$ , as all these representations are related by an bijective affine-linear mapping. The first two cumulants of the summed input  $h_k = \sum_l J_{kl} n_l$  to a neuron are therefore given as before by  $\kappa_{1,k} = \mu_k$  and  $\kappa_{2,k} = \sigma_k^2$  [Eq. (5)].

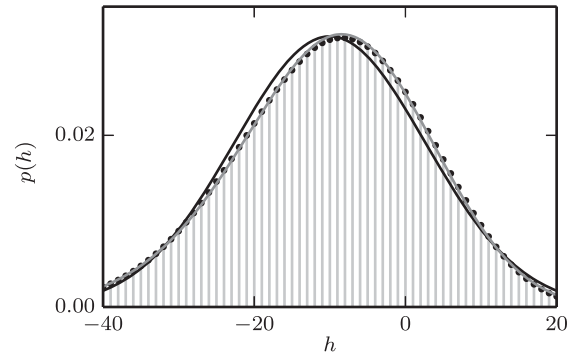


FIG. 3. Distribution of the summed input. Circles show the distribution assuming uncorrelated activity of the input neurons with constant rate  $m = 0.2$ , i.e.,  $p(h) = \sum_{k,l} B(p_{N_E}, m, k) B(p_{N_I}, m, l) \delta_{h, J_{E,k} + J_{I,l}}$ , with binomial probabilities  $B(N, p, k)$ . We show the Gaussian approximation  $\mathcal{N}(\mu, \sigma^2, h)$  (black) with the same moments as the summed binomial distribution  $\mu = p(N_E J_E + N_I J_I) m$ ,  $\sigma^2 = p(N_E J_E^2 + N_I J_I^2) m(1-m)$ . The close-to-Gaussian approximation  $p(h) \approx (1 + (\kappa_3/6) \partial_h^3) \mathcal{N}(\mu, \sigma^2, h)$  (gray) is shown, taking into account the third cumulant  $\kappa_3 = p(N_E J_E^3 + N_I J_I^3) (2m^3 - 3m^2 + m)$  of  $h$ . Other parameters are the same as in Fig. 1.

The third cumulant  $\kappa_{3,k}$  [by defining  $x_l = J_{kl}n_l$  and  $y = h_k$  and using Eq. (B4)] reads

$$\kappa_{3,k} = \sum_{i,j,r} J_{ki} J_{kj} J_{kr} \langle\langle n_i n_j n_r \rangle\rangle, \quad (17)$$

where we use the notation  $\langle\langle \circ \rangle\rangle$  to denote cumulants. For the mean activity (3), we need to evaluate the expectation value of a nonlinear function applied to  $h_k$ . We follow an analogous approach as in the previous section (see Appendix C for details) and apply Eq. (C2) to  $F_k(\mathbf{n}) = H(\sum_j J_{kj} n_j - \theta)$ , yielding a perturbative treatment for the effect of the third-order cumulant

$$\langle F_k(\mathbf{n}) \rangle = e^{(1/6)\kappa_{3,k}(\partial/\partial\mu_k)^3} \langle H(h_k - \theta) \rangle_{h_k \sim \mathcal{N}(\mu_k, \sigma_k^2)}. \quad (18)$$

To obtain a correction of the covariances, we determine the contribution of the third cumulant to the term of the form  $\langle F_k(\mathbf{n}) n_l \rangle = \langle H(h_k(\mathbf{n}) - \theta) n_l \rangle$ . We apply the general result (C3) with  $x_l = J_{kl} n_l$ , where we need to cancel a factor  $J_{kl}$  in the final result because of  $n_l = x_l / J_{kl}$  to get

$$\begin{aligned} \langle F_k(\mathbf{n}) n_l \rangle &= \sum_{q=0}^3 \Delta\kappa_{q,kl} \frac{1}{q!} \left( \frac{\partial}{\partial\mu_k} \right)^q \langle F_k(\mathbf{n}) \rangle \quad \text{with} \\ \Delta\kappa_{q,kl} &= \sum_{i_1 \dots i_q} J_{ki_1} \dots J_{ki_q} \langle\langle n_{i_1} \dots n_{i_q} n_l \rangle\rangle, \end{aligned} \quad (19)$$

where  $\langle F_k(\mathbf{n}) \rangle$  is given by Eq. (18). The form of the terms  $\Delta\kappa_{q,kl}(1/q!)(\partial/\partial\mu_k)^q = \Delta\kappa_{q,kl}(\partial/\partial\mu_k)^q$  shows that, for  $q \geq 1$ , they correspond to an infinitesimal displacement of the  $q$ th cumulant  $\kappa_{q,k}$  by  $\Delta\kappa_{q,kl}$ . These corrections come about by the presence of the variable  $n_l$  in the expectation value, which can alternatively be understood as a conditioned expectation value  $\langle F_k(\mathbf{n}) n_l \rangle = m_l \langle F_k(\mathbf{n}) | n_l = 1 \rangle$ , where the condition  $n_l = 1$  changes the cumulants of  $h_k$ . The first two terms in the sum (19) are identical to the Gaussian approximation in Eq. (12). The difference in the approximation schemes on the level of  $\mathbf{n}$  and  $h_k$ , respectively, is apparent in the term for  $q = 2$ , which contains a correction to the second-order cumulant due to the presence of  $n_l$ . This term is neglected in the Gaussian truncation of cumulants of  $\mathbf{n}$  in Sec. III. The consistent approximation of  $h_k$  up to third-order cumulants analogously requires the term for  $q = 3$ . This correction in turn depends on the fourth-order cumulant  $\langle\langle n_{i_1} \dots n_{i_q} n_l \rangle\rangle$ .

We now use the properties of binary variables that allow us to express a subset of higher cumulants by lower-order ones because of the property  $\langle n_i^K \rangle = \langle n_i \rangle$  (for all integers  $K \geq 1$ ). The  $k$ th raw moment is given by the product of all combinations of lower-order cumulants [60]. For the third moment, this yields

$$\begin{aligned} \langle n_l n_i n_j \rangle &= \langle\langle n_l n_i n_j \rangle\rangle + c_{li} m_j + c_{ij} m_l + c_{jl} m_i \\ &\quad + m_l m_i m_j. \end{aligned} \quad (20)$$

If there are at least two identical indices, we can express the corresponding third-order cumulant by the two lower orders. Both cases  $l = i \neq j$  and  $l = i = j$  lead to the same expression (D1),

$$\langle\langle n_l n_l n_j \rangle\rangle = c_{lj}(1 - 2m_l). \quad (21)$$

In the latter case, we get the third cumulant of a binary variable  $\langle\langle n_l n_l n_l \rangle\rangle = m_l - 3m_l^2 + 2m_l^3$ , which is uniquely determined by its mean. We can therefore take into account the contribution of all third-order cumulants with at least two identical indices to the statistics of  $h_k$ . Stated differently, we calculate the third-order cumulant of  $h_k$  using Eq. (17) and neglecting all cumulants of the binary variables where all indices are different ( $\langle\langle n_i n_j n_l \rangle\rangle \simeq 0$  for  $i \neq j \neq l$ ). A straightforward calculation (see Appendix D Eq. (D2), for details) yields

$$\begin{aligned} \kappa_{3,k} &\simeq [3(\mathbf{J} \otimes \mathbf{J}) \text{diag}(\{1 - 2m_i\}) \mathbf{C} \mathbf{J}^T]_{kk} \\ &\quad - [2(\mathbf{J} \otimes \mathbf{J} \otimes \mathbf{J}) \text{diag}(\{m_i - 3m_i^2 + 2m_i^3\})]_k, \end{aligned} \quad (22)$$

where we use the symbol  $\otimes$  for the element-wise (Hadamard) product of two matrices [61]. For the third cumulant appearing explicitly in Eq. (19), we perform a similar reduction that yields the matrix (D3),

$$\begin{aligned} \Delta\kappa_2 &\simeq \mathbf{J} \otimes \mathbf{J} \text{diag}(\{1 - 2m_i\}) \mathbf{C} \\ &\quad + 2\mathbf{J} \text{diag}(\{1 - 2m_i\}) \otimes (\mathbf{J} \mathbf{C}) \\ &\quad - 2(\mathbf{J} \otimes \mathbf{J}) \text{diag}(\{m_i - 3m_i^2 + 2m_i^3\}). \end{aligned} \quad (23)$$

The matrix  $\Delta\kappa_3$  takes the form (see Appendix D for details)

$$\begin{aligned} \Delta\kappa_3 &\simeq 3\mathbf{J} \otimes ((\mathbf{J} \otimes \mathbf{J}) \{\langle\langle n_i n_j n_l \rangle\rangle_{ij}\}) \\ &\quad + 3(\mathbf{J} \otimes \mathbf{J}) \otimes (\mathbf{J} \{\langle\langle n_i n_j n_l \rangle\rangle_{ij}\}) \\ &\quad + (\mathbf{J} \otimes \mathbf{J} \otimes \mathbf{J}) (\{\langle\langle n_i n_j n_l \rangle\rangle_{ij}\}) \\ &\quad + \text{diag}(\{\langle\langle n_i n_i n_i \rangle\rangle_i\}), \end{aligned} \quad (24)$$

where the fourth-order cumulants are given by Eqs. (D4)–(D6).

To evaluate the expression (18) for the mean activity and Eq. (19) for the covariance, we need the  $n$ th derivative of the complementary error function. We use that  $(d/dx)^{\frac{1}{2}} \text{erfc}(-x) = (1/\sqrt{\pi}) e^{-x^2}$  is a Gaussian and further that the  $n$ th derivative of a Gaussian  $(d/dx)^n e^{-x^2} = (-1)^n H_n(x) e^{-x^2}$  can be expressed in terms of the  $n$ th Hermite polynomial  $H_n$  (see Ref. [62]). Each differentiation with respect to  $\mu_k$  yields an additional factor  $(\sqrt{2}\sigma_k)^{-1}$ . Hence, we get, for  $n \geq 1$ ,

$$\begin{aligned}
L_n(\mu_k, \sigma_k) &:= \left(\frac{d}{d\mu_k}\right)^n \frac{1}{2} \operatorname{erfc}\left(\frac{\theta - \mu_k}{\sqrt{2}\sigma_k}\right) \\
&= \frac{1}{\sqrt{\pi}} (\sqrt{2}\sigma_k)^{-n} H_{n-1}(x) e^{-x^2} \Big|_{x=(\theta-\mu_k)/(\sqrt{2}\sigma_k)}.
\end{aligned}$$

Using this definition and the expansion  $e^{(1/6)\kappa_{3,k}(\partial/\partial\mu_k)^3} \simeq 1 + \frac{1}{6}\kappa_{3,k}(\partial/\partial\mu_k)^3$ , valid for small  $\kappa_{3,k}$  compared to  $\sigma_k^3$ , we get the mean activity

$$\begin{aligned}
m_k = \langle F_k(\mathbf{n}) \rangle &\simeq \left(1 + \frac{1}{6}\kappa_{3,k} \left(\frac{\partial}{\partial\mu_k}\right)^3\right) \frac{1}{2} \operatorname{erfc}\left(\frac{\theta - \mu_k}{\sqrt{2}\sigma_k}\right) \\
&= L_0(\mu_k, \sigma_k) + \frac{1}{6}\kappa_{3,k} L_3(\mu_k, \sigma_k). \quad (25)
\end{aligned}$$

For the term appearing in the covariance, we get

$$\begin{aligned}
\langle F_k(\mathbf{n}) n_l \rangle &\simeq m_l m_k + \sum_{j=1}^N \left( L_1(\mu_k, \sigma_k) + \frac{1}{6}\kappa_{3,k} L_4(\mu_k, \sigma_k) \right) J_{kj} c_{jl} \\
&\quad + \frac{1}{2} \left( L_2(\mu_k, \sigma_k) + \frac{1}{6}\kappa_{3,k} L_5(\mu_k, \sigma_k) \right) \Delta\kappa_{2,kl} \\
&\quad + \frac{1}{6} \left( L_3(\mu_k, \sigma_k) + \frac{1}{6}\kappa_{3,k} L_6(\mu_k, \sigma_k) \right) \Delta\kappa_{3,kl} + o(\kappa_{3,k}^2). \quad (26)
\end{aligned}$$

The latter expression shows, in particular, that the correction to the mean activity reappears (as the factor  $m_k$ ) in the first line, indicating that the contribution of the third cumulant to lowest order [ $q = 0$  in Eq. (19)] drops out of the covariance, as the term  $m_l m_k$  is subtracted in Eq. (3). In a weakly correlated state of the network, the remaining terms are smaller than  $m_l m_k$  as they give rise to the pairwise covariance (3). This explains why the Gaussian approximation is already fairly accurate.

Simultaneously solving Eqs. (25) and (26) by a damped fixed-point iteration, analogous to Eq. (16), yields an approximation of the mean activities and covariances shown in Fig. 4. The deviation of the mean activities from simulation results [Fig. 4(a)] is reduced compared to the Gaussian approximation below the significance level. The width of the distributions is only slightly underestimated compared to simulations, as exhibited by the scatter plot aligned with the diagonal and the bar graph. The pairwise averaged cross-covariances [Fig. 4(b)] are within the error of the simulated results, in contrast to the Gaussian approximation [cf. Fig. 1(b)]. A small contribution to the remaining difference in variance stems from the finite simulation time, naturally leading to a wider distribution in Fig. 4(b) compared to the theoretical prediction. The

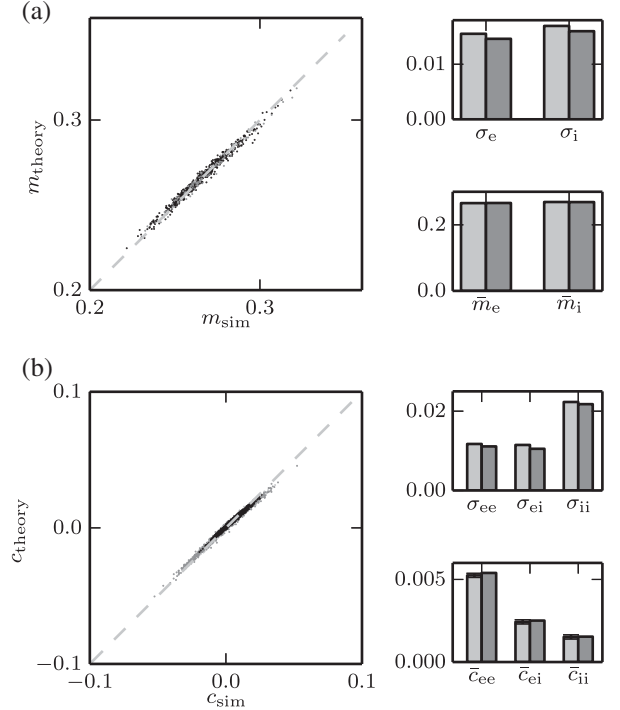


FIG. 4. Approximation for covariance matrix in a random network of binary neurons including third-order cumulants. (a) Theoretical prediction of mean activities (25) versus simulation results (black is for excitatory; gray is for inhibitory). Lower inset: Mean activity averaged over excitatory and inhibitory neurons (light gray is for simulation; darker gray is for theory). Black error bars (not visible, below line width) show the standard error of the mean obtained from 20 simulations. Upper inset: Standard deviation of distribution of mean activities of excitatory ( $\sigma_e$ ) and inhibitory ( $\sigma_i$ ) neurons. (b) Theoretical prediction (26) versus simulated covariance (black is for excitation-excitation; gray is for inhibition-inhibition; light gray is for excitation-inhibition). Lower inset: Mean covariances averaged over excitatory-excitatory, excitatory-inhibitory, and over inhibitory-inhibitory neuron pairs (light gray is for simulation; darker gray is for theory). Error bars show the standard error of the mean obtained from 20 simulations. Upper inset: Standard deviations of the distributions of covariances. Network parameters and the display are the same as in Fig. 1. Theoretical predictions are obtained by damped fixed-point iteration.

bigger contribution presumably comes from the nontrivial third-order cumulants  $\langle\langle n_i n_j n_k \rangle\rangle$ , where all indices are unequal. Neglecting the nontrivial covariances shows that the variability of  $\sigma_k$  from neuron to neuron is reduced, which in turn reduces the width of the distribution of the mean activities. For networks with fixed in-degree, this even leads to a uniform mean activity across neurons, which, however, still matches the averaged mean activities from simulations indicating that averaged quantities are insensitive to variability in nontrivial higher-order cumulants (see Fig. 5 in Sec. VI). Analogously, we expect that the neglect of nontrivial third-order cumulants underlies the reduced width of the covariances.



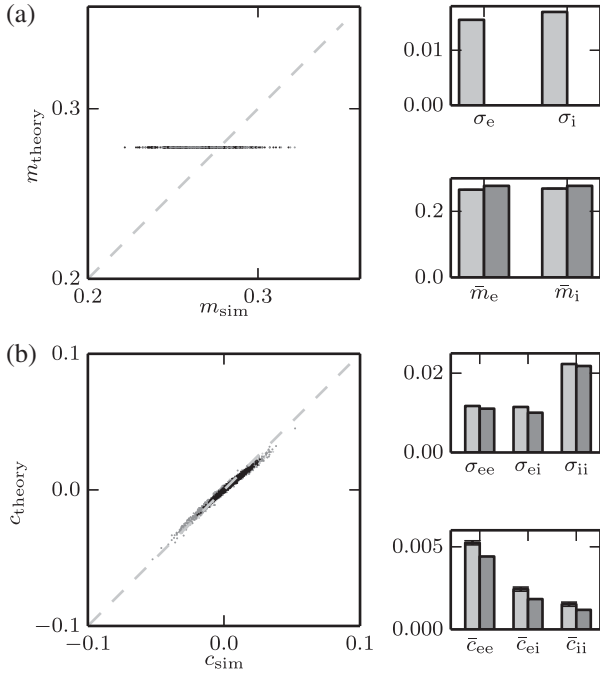


FIG. 5. Modified Lyapunov method. Solution of the individual pairwise covariances by the projection method applied to the continuous Lyapunov equation in comparison to simulation results. (a) Mean activities predicted by the approximation that neglects cross-correlations (black is for excitatory; gray is for inhibitory). Lower inset: Mean activity averaged over excitatory and inhibitory neurons (light gray is for simulation; darker gray is for theory). Black error bars show the standard error of the mean obtained from 20 simulations (barely visible, approximately at line width). Upper inset: Width of distribution of mean activities. (b) Theoretical distribution of covariances (29) with  $\mathbf{D}$  chosen according to Eq. (30) [same approximation as in (a) for  $\mathbf{m}$  and  $\mathbf{S}$ , i.e., neglecting cross-correlations in the input to each neuron]. Lower inset: Mean covariances averaged over excitatory-excitatory, excitatory-inhibitory, and over inhibitory-inhibitory neuron pairs (light gray is for simulation; darker gray is for theory). Error bars show the standard error of the mean obtained from 20 simulations. Upper inset: Width of the distributions of covariances. Other parameters and the display are the same as in Fig. 1.

## V. SCALE INVARIANCE OF COVARIANCES

The equation for cross-covariances in Eq. (15) yields an additional insight: Assuming that mean activities and correlations were unchanged, a scaling of all incoming synapses to a neuron  $k$  by some factor  $\alpha > 0$  amounts to a scaling of the strength of synaptic fluctuations in the same manner ( $\sigma_k \propto \alpha$ ). The fixation of mean activities can be achieved for different choices of incoming synaptic amplitudes by adapting the threshold such that  $(\mu_k(\{J_{kl}\}) - \theta_k)/\sigma_k$  remains invariant [cf. Eq. (15)]. A uniform rescaling of all synapses by a factor  $\alpha > 0$  therefore requires a change of the threshold by the same factor. The susceptibility [Eq. (14)]  $S_k$  then scales as  $\sigma_k^{-1} \propto \alpha^{-1}$ . Hence, the term  $S_k J_{kl}$  appearing in the equation for

covariances in Eq. (15) is invariant under this scaling. This result implies that covariances are invariant with respect to the absolute value of synaptic amplitudes: The self-generated network noise causes a divisive normalization on the level of the synaptic input to each neuron. This also implies that scaling the synapses with  $J \propto N^{-1}$ ,  $J \propto N^{-(1/2)}$ , or  $J \propto 1$  as the network size tends to infinity yields the same covariances, given the thresholds are adapted so that the mean activity is preserved. On the level of population-averaged covariances, this idea has been noted earlier [59]. The independence of the network dynamics on the absolute value of the synaptic weights even holds exactly, as noted earlier [63]. The reason for this is the absence of a length scale of a hard threshold; the only relevant length scale is the amplitude  $\sigma_k$  of the synaptic noise itself. Considering a single neuron  $k$ , the condition for the neuron to be activated is  $h_k = \sum_l J_{kl} n_l > \theta_k$ . Rescaling all incoming synapses as well as the threshold by the same factor  $\alpha > 0$  multiplies both sides of this inequality; the neuron switches at the very same configuration  $\mathbf{n}$  of incoming spins as in the original case.

The previous consideration is completely in line with the work of Ref. [56]. The latter work found that, when increasing the number of neurons  $N$ , a scaling of  $1/\sqrt{N}$  is required to obtain a robust asynchronous state, if the system possesses variability of the thresholds with a standard deviation defined as unity: The width of the distribution of the thresholds induces a second length scale into the system. Invariant behavior can therefore only arise if the synaptic noise  $\sigma_k$  and the standard deviation of the thresholds have a constant ratio. Scaling synapses as  $J \propto 1/N$  in this setting leads to vanishing temporal fluctuations of  $h_k$  and hence spin-glass freezing, whereas for  $J \propto 1$  the asynchronous state persists. Fixing  $J$ , while increasing the number of neurons  $N$ , results in more inputs per neuron and thus increased temporal fluctuations, which wash out the effect of distributed thresholds.

## VI. MAPPING OF FLUCTUATIONS TO ORNSTEIN-UHLENBECK PROCESSES

Here, we show an alternative interpretation of the Gaussian approximation in Sec. III. Despite the fact that the covariances obey different equations on the diagonal  $k = l$  and the off-diagonal  $k \neq l$  [Eq. (15)], we can rewrite the equations as a single matrix equation

$$2\mathbf{C} - \mathbf{D} = \mathbf{WC} + (\mathbf{WC})^T$$

$$0 = (\mathbf{W} - \mathbf{1})\mathbf{C} + ((\mathbf{W} - \mathbf{1})\mathbf{C})^T + \mathbf{D}, \quad (27)$$

where  $\mathbf{D}$  is a diagonal matrix with elements constrained by the condition that the diagonal entries of the covariance matrix fulfill [Eq. (3)]

$$c_{kk} = m_k(1 - m_k), \quad (28)$$

and  $\mathbf{W} = \mathbf{S}\mathbf{J}$  is the matrix of effective couplings, which incorporates the susceptibility  $\mathbf{S} = \text{diag}(S_1, \dots, S_N)$  of the individual units. We ensure this latter condition in the following way. We first solve Eq. (27) by multiplication with the left eigenvectors  $\mathbf{v}_\alpha$  of the effective couplings, i.e.,  $\mathbf{v}_\alpha^T(\mathbf{W} - \mathbf{1}) = \lambda_\alpha \mathbf{v}_\alpha^T$  (see Chap. 6.5 of Ref. [46]). The corresponding right eigenvectors are  $\mathbf{u}_\alpha$ , which fulfill the bi-orthogonality  $\mathbf{v}_\alpha^T \mathbf{u}_\beta = \delta_{\alpha\beta}$  and completeness  $\sum_\alpha \mathbf{u}_\alpha \mathbf{v}_\alpha^T = \mathbf{1}$  relation. With the notation  $C^{\alpha\beta} := \mathbf{v}_\alpha^T \mathbf{C} \mathbf{v}_\beta$  (analogously for  $\mathbf{D}$ ), we have

$$\begin{aligned} C^{\alpha\beta} &= -\frac{D^{\alpha\beta}}{\lambda_\alpha + \lambda_\beta}, \\ \mathbf{C} &= \sum_{\alpha,\beta} \mathbf{u}_\alpha \mathbf{u}_\beta^T C^{\alpha\beta}, \end{aligned} \quad (29)$$

where the latter expression follows from the completeness relation.

The expression reveals the connection between the eigenvalues  $\lambda_\alpha$  of the linearized coupling matrix and the fluctuations in the system. In the case of a single eigenvalue close to instability, i.e.,  $\text{Re}(\lambda_\alpha) \approx 0$ , the fluctuations in the corresponding eigendirection constitute the dominant contribution to the covariance matrix  $\mathbf{C} \approx -\mathbf{u}_\alpha \mathbf{u}_\alpha^T (\mathbf{v}_\alpha^T \mathbf{D} \mathbf{v}_\alpha) / (2\lambda_\alpha)$ . This scenario could be detected in a principle-component analysis of experimental data, with the dominant principle component pointing in the direction of  $\mathbf{u}_\alpha$ .

Evaluating expression (29) on the diagonal and using  $D^{\alpha\beta} = \mathbf{v}_\alpha^T \mathbf{D} \mathbf{v}_\beta = \sum_j v_{\alpha,j} v_{\beta,j} D_j$ , we get

$$\begin{aligned} m_k(1 - m_k) \stackrel{!}{=} c_{kk} &= \sum_j - \underbrace{\left( \sum_{\alpha,\beta} \frac{u_{\alpha,k} v_{\alpha,j} \cdot u_{\beta,k} v_{\beta,j}}{\lambda_\alpha + \lambda_\beta} \right)}_{\equiv B_{kj}} D_j, \\ \mathbf{B} &= - \sum_{\alpha,\beta} \frac{(\mathbf{u}_\alpha \mathbf{v}_\alpha^T) \otimes (\mathbf{u}_\beta \mathbf{v}_\beta^T)}{\lambda_\alpha + \lambda_\beta}, \end{aligned} \quad (30)$$

where the symbol  $\otimes$  is to be understood as the element-wise multiplication (Hadamard product) of the two matrices in the numerator. The penultimate line is an ordinary matrix equation relating the ( $N$ -dimensional vector)  $D_k$  to the ( $N$ -dimensional vector)  $c_{kk}$ . To determine  $\mathbf{D}$  as  $\mathbf{D} = \mathbf{B}^{-1} \text{diag}(\{c_{kk}\})$ , we hence need to invert the matrix  $\mathbf{B}$ . The covariance matrix is then obtained by Eq. (29).

The result (27) can be further interpreted as a mapping of the fluctuations from the binary dynamics to an effective system of Ornstein-Uhlenbeck processes. Given the set of coupled Ornstein-Uhlenbeck processes

$$\tau \frac{\partial x_k(t)}{\partial t} = -x_k(t) + \sum_j w_{kj} x_j(t) + \xi_k(t), \quad (31)$$

with the Gaussian white noise  $\langle \xi(t) \xi^T(s) \rangle = \tau \mathbf{D} \delta(t - s)$ , the stationary equal-time covariance matrix fulfills the

same continuous Lyapunov equation (see Chap. 6.5 of Ref. [46]) as the binary network (27). Using the analogy of the continuous Lyapunov equations, we see that the elements of the diagonal matrix  $\mathbf{D}$  can be interpreted as the noise intensity injected into each neuron. The modified Lyapunov method (30) determines this intensity such that the variance of each continuous variable  $x_k$  agrees with the variance of the corresponding binary variable  $n_k$  given by Eq. (28), which, in turn, is fixed by the mean activity (15).

It is important to note that  $\mathbf{W}$  and  $\mathbf{D}$  in Eqs. (27) and (31) themselves depend on the covariances  $\mathbf{C}$  via the susceptibility  $\mathbf{S}$ . This leads to Eq. (29) remaining an implicit equation for  $\mathbf{C}$  that needs to be solved iteratively. However, by neglecting the contribution of cross-covariances in Eq. (5),  $\mathbf{W}$  and  $\mathbf{D}$  become independent of  $\mathbf{C}$ , rendering Eq. (27) a linear equation for the covariances and the above projection method an efficient algorithm to compute  $\mathbf{C}$ . This linear approximation consequently fails to predict the distribution of mean activities for networks with fixed in-degree, as shown in Fig. 5(a). Nevertheless, the width of the distribution of the covariances shown in Fig. 5(b) is only slightly underestimated, showing that the distribution of mean activities contributes only marginally to the distribution of covariances.

The viability of the linear approximation for the covariances shows that fluctuations in the binary model are practically equivalent to those in linear Ornstein-Uhlenbeck processes. This equivalence has been reported earlier for effective equations of the population-averaged pairwise correlations [54]. The latter work used the approximate result  $c_{kk} \approx (D_k/2)$ , which ignores the effect of the covariances onto the autocovariances in Eq. (27), i.e., assuming that  $(\mathbf{W}\mathbf{C})_{kk}$  is smaller than  $c_{kk}$  itself and hence neglecting the right-hand side of Eq. (27) when determining the diagonal elements  $c_{kk}$ . In the absence of self-connections, this amounts to neglecting the off-diagonal cross-covariances in  $\mathbf{C}$ . For weakly correlated network states, this is a good approximation. However, in the present network setting, it slightly overestimates the population averages of the covariances as well as the width of their distribution (data not shown).

## VII. RESPONSE OF THE NETWORK TO EXTERNAL STIMULI

To study how the recurrent network processes an externally applied signal, we take into account an additional external input to each neuron  $k$ , denoted as  $h_{k,\text{ext}}(t)$ , such that  $h_k(t) = \sum_j J_{kj} n_j(t) + h_{k,\text{ext}}(t)$ . We start from the differential equation (2) for the mean activities

$$\tau \frac{\partial}{\partial t} \langle n_k(t) \rangle = -\langle n_k(t) \rangle + \langle F_k(\mathbf{n}(t), h_{k,\text{ext}}(t)) \rangle,$$

with

$$F_k(\mathbf{n}(t), h_{k,\text{ext}}(t)) = H\left(\sum_j J_{kj} n_j(t) + h_{k,\text{ext}}(t) - \theta\right).$$

As in the case without external input, the expectation value  $\langle F_k(\mathbf{n}(t), h_{k,\text{ext}}(t)) \rangle$  can be treated in the Gaussian approximation. Subtracting the stationary activity state  $\delta n_k(t) = n_k(t) - \langle n_k \rangle$ , we get, after linearization,

$$\begin{aligned} \tau \frac{\partial}{\partial t} \langle \delta n_k(t) \rangle &= -\langle \delta n_k(t) \rangle + S_k \left( \sum_j J_{kj} \langle \delta n_j(t) \rangle + h_{k,\text{ext}}(t) \right) \\ &= \sum_j (w_{kj} - \delta_{kj}) \langle \delta n_j(t) \rangle + \underbrace{S_k h_{k,\text{ext}}(t)}_{=: y_k(t)}, \end{aligned}$$

where  $w_{kj} = S_k J_{kj}$  is the effective coupling as defined in Sec. VI and  $S_k$  is the susceptibility (14). Hence, the equation of motion for the perturbation in matrix notation reads

$$\tau \frac{\partial}{\partial t} \langle \delta \mathbf{n}(t) \rangle = (\mathbf{W} - \mathbf{1}) \langle \delta \mathbf{n}(t) \rangle + \mathbf{y}(t). \quad (32)$$

In order to excite the system into the direction of one eigenmode of the effective couplings  $\mathbf{W}$ , we choose the following stimulus vector:

$$\mathbf{h}_{\text{ext}}^\alpha(t) := a \tilde{\mathbf{u}}_\alpha f(t), \quad \text{with} \quad \tilde{\mathbf{u}}_\alpha = \frac{\mathbf{S}^{-1} \text{Re}(\mathbf{u}_\alpha)}{\|\mathbf{S}^{-1} \text{Re}(\mathbf{u}_\alpha)\|}, \quad (33)$$

where  $\mathbf{S} := \text{diag}(\{S_k\})$  is the diagonal matrix containing the susceptibilities and  $\mathbf{u}_\alpha$  the right-sided eigenvector of  $\mathbf{W} - \mathbf{1}$  as defined in Sec. VI. The strength of the stimulus is regulated by the parameter  $a$  and its temporal profile determined by  $f(t)$ . The parameter  $a$  allows us to control the amplitude of the stimulation in comparison to the strength of the synaptic noise received by each unit. For the linear approximation to hold, this amplitude must be chosen such that the input to each unit is in the linear part of the expectation value of the gain function and can hence be approximated by the slope (14). Stimulating into the direction of the real part of the eigenvectors ensures that a complex mode is excited in combination with its complex conjugate, which is necessary since the activity in the network is real-valued. Here,  $\tilde{\mathbf{u}}_\alpha$  is constructed such that it is normalized and compensates for the multiplication of the external input with the diagonal susceptibility matrix. We measure the deflection of the activity into the direction of the eigenmode by defining  $\langle \delta n^\alpha(t) \rangle := \tilde{\mathbf{v}}_\alpha^T \langle \delta \mathbf{n}(t) \rangle$  as the projection of the activity vector onto the  $\alpha$ th eigenmode of the effective coupling matrix  $\mathbf{W}$  ( $y^\alpha$  defined analogously), where we choose  $\tilde{\mathbf{v}}_\alpha^T$  such that  $\tilde{\mathbf{v}}_\alpha^T \mathbf{S} \tilde{\mathbf{u}}_\beta = \delta_{\alpha\beta}$ , i.e.,

$$\begin{aligned} \tilde{\mathbf{v}}_\alpha &:= k \|\mathbf{S}^{-1} \text{Re}(\mathbf{u}_\alpha)\| \text{Re}(\mathbf{v}_\alpha), \quad \text{with} \\ k &= \begin{cases} 1 & \text{if } \text{Im}(\lambda_\alpha) = 0 \\ 2 & \text{if } \text{Im}(\lambda_\alpha) \neq 0, \end{cases} \end{aligned}$$

with  $\mathbf{v}_\alpha$  being the left eigenvectors of  $\mathbf{W} - \mathbf{1}$  as defined in Sec. VI.

The time evolution of the perturbation  $\langle \delta n^\alpha(t) \rangle$  is obtained by solving Eq. (32) projected onto  $\mathbf{v}_\alpha$  and  $\mathbf{v}_\alpha^*$ , and subsequently adding the results. Hence, for any stimulus  $f(t)$  [inserted into Eq. (32) as  $\mathbf{y}(t) = \mathbf{S} \mathbf{h}_{\text{ext}}^\alpha(t)$ ], the perturbation obeys the convolution equation

$$\langle \delta n^\alpha \rangle(t) = a \frac{1}{2\tau} [(e^{\lambda_\alpha(\circ/\tau)} + e^{\lambda_\alpha^*(\circ/\tau)}) * f](t). \quad (34)$$

Here, we consider a piecewise constant input starting at time  $t_0 = 0$  and stopping at  $t = T$ ; i.e., we choose  $f(t) = H(t) - H(t - T)$ . The time course of the perturbation is then given by

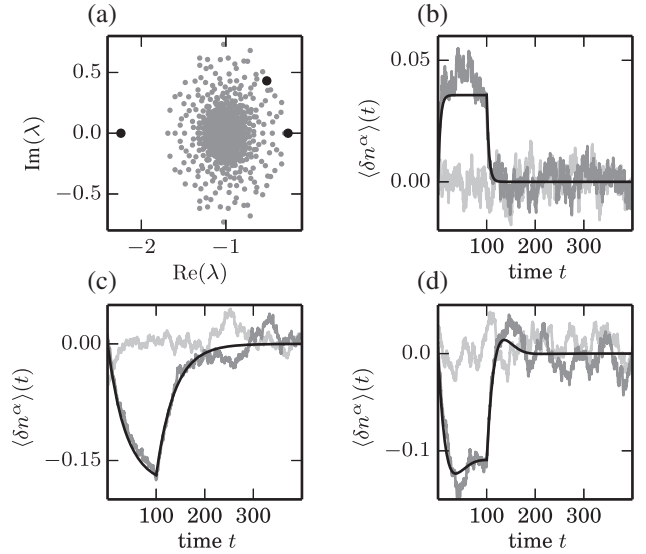


FIG. 6. Linear responses to DC stimuli. Responses (normalized by the number of units  $N$ ) to a DC stimulus (duration  $T = 100$  ms) in the direction of one eigenmode (dark gray), the background activity (light gray) measured in simulation, and the analytical prediction of the response curve [Eq. (35), black] obtained from linear response theory. (a) Eigenvalues  $\lambda_\alpha$  of the matrix  $\mathbf{W} - \mathbf{1}$ . The eigenvalues associated with the modes that are stimulated are depicted by black dots. (b) Stimulation in the direction of the fastest decaying eigenmode  $\lambda = -2.24$ , with a stimulus size  $a = 50$ . (c) Stimulation in the direction of the slowest decaying eigenmode  $\lambda = -0.26$ , with a stimulus size  $a = -30$ . (d) Stimulation in the direction of an eigenmode decaying slowly and oscillatory  $\lambda = -0.52 + 0.43i$ , with a stimulus size  $a = -60$ . Responses averaged over 500 repetitions. Other parameters are the same as in Fig. 1.

$$\langle \delta n^a \rangle(t) = a \begin{cases} \text{Re}(\lambda_a^{-1}(1 - e^{\lambda_a t/\tau})) & \text{if } t \leq T \\ \text{Re}(\lambda_a^{-1}(e^{\lambda_a(t-T)/\tau} - e^{\lambda_a t/\tau})) & \text{if } t > T. \end{cases} \quad (35)$$

Figure 6 shows the response of the network activity projected onto one eigenmode to the stimulus in the direction of the same eigenmode. Choosing three stimulus directions associated with three representative eigenvalues from the full eigenvalue cloud of the network [Fig. 6(a)], we demonstrate that the time course of fast (large negative eigenvalue), slowly (eigenvalue close to zero), and oscillatory (complex eigenvalue) decaying modes is captured by the linear approximation. This shows how the response of the network depends on the spatial structure of the stimulation. To lowest order, the transformation performed by the network is hence a spatiotemporal filtering of the input, where the responses with slowest decay correspond to the excitation of eigenmodes closest to instability.

### VIII. RECONSTRUCTION OF COUPLINGS

In the current section, we investigate the inverse problem, i.e., the inference of the couplings  $J_{kl}$  of the network from the observed activity. The time-lagged cross-covariance function in a network of Ornstein-Uhlenbeck processes (for  $t > s$ ) fulfills the differential equation (see Chap. 6.5 of Ref. [46])

$$q_{kl} := \tau \frac{\partial}{\partial t} c_{kl}(t, s) + c_{kl}(t, s) \stackrel{t \geq s}{=} \sum_j w_{kj} c_{jl}(t, s). \quad (36)$$

This means we can uniquely reconstruct the effective couplings  $w_{kl}$  from the knowledge of the two matrices, the covariance and its slope at time lag zero as

$$\begin{aligned} \mathbf{W} &= \mathbf{Q}\mathbf{C}^{-1} \\ &= \mathbf{1} + \tau \frac{\partial}{\partial t} \mathbf{C}(t, s)|_{t=s} \mathbf{C}^{-1}. \end{aligned} \quad (37)$$

For an Ornstein-Uhlenbeck process, obeying Eqs. (27) and (36), the reconstruction of the couplings  $w_{kl}$  is exact, as shown in Fig. 8(b).

We now demonstrate that a similar relationship approximately also holds in binary networks. For  $t > s$ , the time-lagged cross-covariance function  $c_{kl}(t, s) = \langle n_k(t)n_l(s) \rangle - \langle n_k(t) \rangle \langle n_l(s) \rangle$  for the binary network fulfills [Eq. (A5)]

$$\begin{aligned} q_{kl}(t, s) &:= \tau \frac{\partial}{\partial t} c_{kl}(t, s) + c_{kl}(t, s) \\ &\stackrel{t \geq s}{=} \langle F_k(\mathbf{n}(t))n_l(s) \rangle - \langle n_k(t) \rangle \langle n_l(s) \rangle. \end{aligned} \quad (38)$$

An example of an autocovariance and a cross-covariance function together with the slope at zero time lag is shown in Fig. 7.

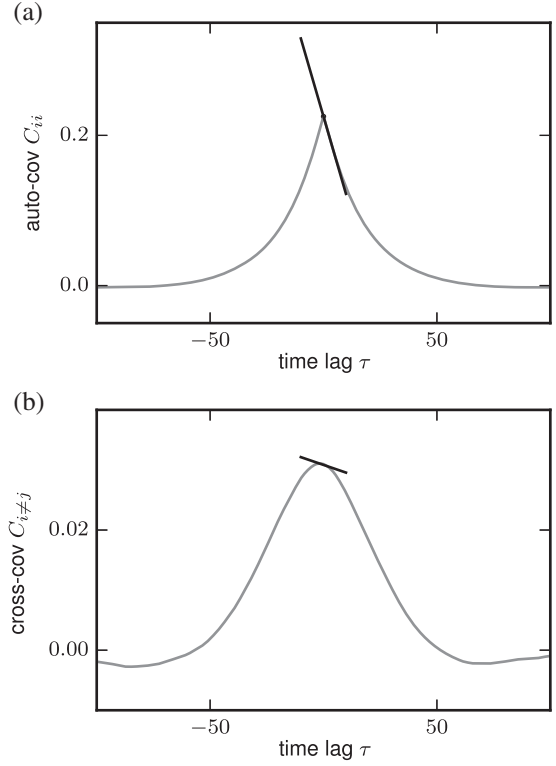


FIG. 7. Autocovariance and cross-covariance functions. Example of an autocovariance function (a) and a cross-covariance function (b) in the network described in Fig. 1. The respective slopes of the functions at time lag 0+ are indicated by the black tangent lines.

In the limit of vanishing time lag, the form of  $q_{kl} = \lim_{t \rightarrow s} q_{kl}(t, s) = \langle F_k(\mathbf{n}(t))n_l(t) \rangle - \langle n_k(t) \rangle \langle n_l(t) \rangle$  shows that  $q_{kl}$  measures the direct influence of fluctuations of neuron  $l$  on the gain function of neuron  $k$ . In the Gaussian approximation (see Sec. III), we get for  $k \neq l$

$$q_{kl} = \sum_j s_{kj} J_{kj} c_{jl}. \quad (39)$$

For  $k = l$ , however, we have to evaluate  $\langle H(h_k(\mathbf{n}(t) - \theta)n_k(t)) \rangle$  on the right-hand side. Since the term  $n_k$  appears in the expectation value, the statistics of  $h_k$  is effectively conditioned on the state  $n_k = 1$ . Since the transitions of neuron  $k$  directly depend on the value of  $h_k$ , this conditioning violates the close-to Gaussian approximation of  $h_k|n_k = 1$ . An approximate treatment on the diagonal is possible under the assumption that the autocovariance function of  $h_k$  is dominated by contributions of the autocovariances of the binary variables  $n$ , yielding a nonlinear differential equation for the temporal shape of the autocovariance function [see Eq. (5.17) of Ref. [56]]. Approximating the temporal profile of the autocovariance function  $a(s) := \langle H(h_k(t+s) - \theta)H(h_k(t) - \theta) \rangle \approx \langle m_k \rangle + \langle m_k \rangle(1 - \langle m_k \rangle)e^{-|s|/\tau}$ , which has the right asymptotic



behaviors [ $a(0) = m_k$  and  $a(s \rightarrow \infty) = m_k^2$ ] and assumes fluctuations to be correlated on the time scale  $\tau$  of the update, leads to the approximate value  $\langle H(h_k(t) - \theta)n_k(t) \rangle \simeq \frac{1}{2}m_k(1 - m_k)$  [by using Eq. (C1) of Ref. [56]]. The corresponding approximation of the slope following from the former approximation with Eq. (38) is shown in Fig. 8(a). However, the resulting expression for the slope at zero time lag cannot be written in the form (39). Therefore, we assume that Eq. (39) also holds on the diagonal, although the resulting predicted slopes have a slight negative bias compared to simulation results.

Given the two covariance matrices entering Eq. (38), as well as the average activities  $m_i$  for each neuron, we can derive a procedure for reconstructing the couplings  $J_{kl}/\sigma_k$  relative to the noise. The mean and variance of the input to a neuron are determined by Eq. (5). Inverting the relation (6)

as  $y(m_k) := (\mu_k - \theta)/(\sqrt{2}\sigma_k) = -\text{erfc}^{-1}(2m_k)$  allows us to determine the susceptibility (14) as  $S(m_k, \sigma_k) = 1/(\sqrt{2\pi}\sigma_k)e^{-y(m_k)^2}$ . Hence, we obtain an expression for the ratio of the incoming synaptic weight and the total synaptic noise of neuron  $k$ ,

$$\frac{J_{kl}}{\sigma_k} = \sqrt{2\pi}w_{kl}e^{y(m_k)^2}. \quad (40)$$

This result again shows that correlations are only controlled by the ratio  $J_{kl}/\sigma_k$  rather than by  $J_{kl}$  and  $\sigma_k$  alone, in line with the invariance found in Sec. V. Figure 8(b) shows the reconstruction of couplings from the simulated covariance functions. Because of the previously discussed approximations, the exact reconstruction of the relative couplings  $J_{kl}/\sigma_k$  is not possible for a binary network. The observed deviations from the identity line are predominantly caused by the approximation of the slope of the autocovariance functions. The reconstructed coupling matrix correctly infers all excitatory and all inhibitory connections but additionally yields a considerable number of false-positive excitatory and inhibitory connections.

The described procedure, moreover, allows us to determine the remaining parameters of the binary network. With regard to correlations, we are free to choose an arbitrary  $\sigma_k$ , e.g.,  $\sigma_k = 1$  to determine  $J_{kl}$  by Eq. (40). As the mean activity and correlations are known, the actual magnitude of fluctuations  $\sigma_k^{\text{loc}}$  caused by the local inputs from the network follows from Eq. (5). If these fluctuations are smaller than or equal to our arbitrary choice ( $\sigma_k^{\text{loc}} \leq \sigma_k = 1$ ), it is possible to supply each neuron with additional noise of variance  $1 - (\sigma_k^{\text{loc}})^2$ . Only in this case can we construct a binary network satisfying the given constraints. Having fixed  $J_{kl}$ , and given the mean activities by Eq. (5), we can determine the mean input  $\mu_k$  to each cell. Finally, the threshold  $\theta$  must be chosen such that  $\theta_k = \mu_k - \sqrt{2}y(m_k)$ .

Different routes to inverse problems have been proposed earlier. For equilibrium systems, a statistical-mechanics formulation allows the use of principal-component analysis [64] or maximum-entropy approaches [65,66]. For off-equilibrium systems with a sequential Glauber update in discrete time steps, methods related to our approach have previously been proposed by Mézard and Sakellariou [42]. Their results can be related to our continuous-time method by time discretization of the slope appearing on the left-hand side of Eq. (38), leading to similar expressions as their Eqs. (8) and (18). An alternative approach, based on the Thouless-Anderson-Palmer (TAP) approximation [5], is described in Ref. [41] [see their Eq. (14)].

## IX. DISCUSSION

Here, we present a theoretical description of fluctuations in strongly coupled networks of large numbers of binary units that, for the first time, faithfully captures the statistics

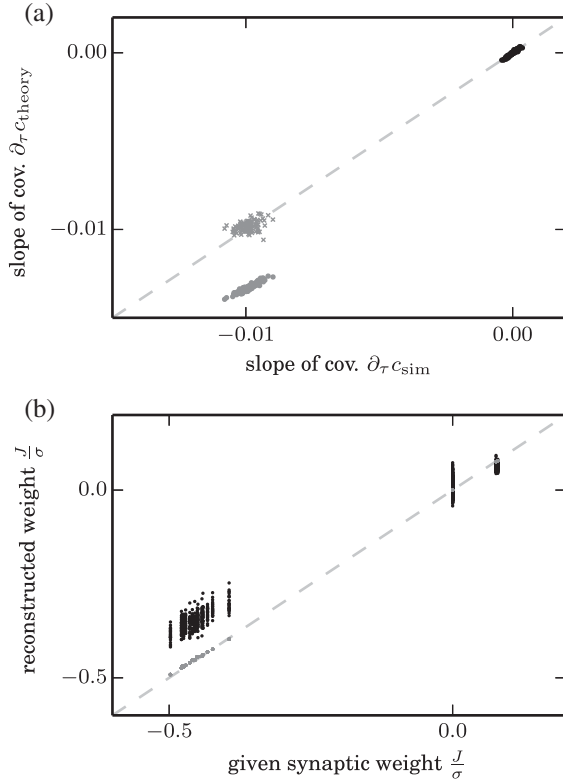


FIG. 8. Reconstruction of couplings from the matrix of covariances and slopes of the covariance functions at zero time lag. (a) Slope of autocovariance (gray) and cross-covariance functions (black) at zero time lag from the Gaussian approximation versus slope taken from simulation results. Dots show the prediction of the linear theory (36), and crosses the approximation  $-\frac{1}{2}m_k(1 - m_k)$  for the slope of autocovariances. (b) Reconstructed synaptic amplitude  $w_{kl}$  in a network of Ornstein-Uhlenbeck processes (37) (gray) and reconstructed weights  $J_{kl}/\sigma_k$  (40) (black) in a binary network versus original couplings used in simulation. Covariance and slope averaged over 20 repetitions, each simulated for  $T = 2,000,000$  ms. Other network parameters are the same as in Fig. 1.

of individual units as well as pairs of units. The fluctuations are characterized by self-consistent equations for the mean activity and pairwise correlations, including finite-size effects down to hundreds of units. The method can be applied to a wide range of networks, in particular, networks with asymmetric and strong couplings.

Standard approaches describing symmetric systems in terms of a partition function cannot be extended to non-symmetric coupling because these systems do not reach thermodynamic equilibrium. Asymmetries arising from disordered couplings or deterministic constraints, such as Dale’s law, are treated by using ensemble averages over the random couplings [67] or the random dilution of the network [68], respectively. By construction, these approaches are, however, constrained to describe self-averaging properties of systems. Here, we are interested in the dynamics of individual units, which is not self-averaging, and we set out to capture the statistics of a particular realization of the system in terms of the cumulant hierarchy for the activity variables. Truncating the hierarchy after second order yields a closed set of equations that already provides a good approximation for the zero time lag covariances between individual units. We show the equivalence of this truncation to the Gaussian approximation of the input field [34,56], which follows from the central-limit theorem. Non-Gaussian, finite-size effects in networks of only several hundreds of units are effectively taken into account by incorporating a subset of third-order cumulants that can be expressed in terms of lower-order cumulants due to the binary nature of the activity variables. The inclusion of third-order cumulants significantly improves the prediction of mean activities, which are strongly affected by the nonlinearity of the gain function. The resulting set of nonlinear coupled equations can efficiently be solved numerically by damped fixed-point iteration.

We demonstrate that a Heaviside activation function— independent of the choice of coupling amplitudes— constitutes the strongest possible interaction for binary networks. This finding generalizes the invariance of population-averaged pairwise correlations under proportional scaling of couplings and threshold of all units [59,63] to the invariance of individual pairwise correlations under the scaling of the corresponding units’ parameters. Weaker coupling in binary networks can only be achieved by a smoother activation function, not by different scalings of coupling amplitudes, e.g.,  $1/N$  versus  $1/\sqrt{N}$ , contradictory to frequently employed arguments in the literature (as discussed in Ref. [63]). In networks with plastic synapses or networks receiving time-varying input, dynamically changing thresholds could similarly act as a homeostasis mechanism on the level of mean activities and correlations and prevent systems from freezing in a local minimum [69]. While the scaling invariance is generic, the presented Gaussian and close-to-Gaussian approximations of the input field hold for the commonly considered narrowly

distributed couplings. For wide distributions, the statistics of the input field is likely to depart from the close-to-Gaussian assumption.

The contribution of cross-covariances to the marginal statistics of the input to each neuron causes a distribution of the mean activities, even in networks composed of neurons each receiving an identical number of inputs. The classical treatment of neurons with a Heaviside non-linearity neglects this effect [56]. Still, distributed numbers of synaptic input typically dominate the distribution of mean activities in sparsely [56] and densely connected random networks [34,55].

The Gaussian closure, i.e., the truncation of the cumulant hierarchy after second order, yields a set of equations that accurately predicts the second-order statistics. This result is in line with experimental evidence, showing that pairwise correlations sufficiently constrain maximum-entropy models of collective activity [70]. Analogously, the truncation after first order, i.e., the Curie-Weiss mean-field theory, neglects fluctuations of the local order parameter. Formally, this theory can be obtained as the leading order of a saddle-point approximation in the auxiliary fields (see, e.g., Sec. 4.3 of Ref. [71]). Still, it yields a good estimate of the first moments. In the context of balanced networks, this description is sometimes referred to as “balance equations” [see Eqs. (4.1) and (4.2) of Ref. [56]]. Taking into account fluctuations in the input, due to the variance of individual units [56], constitutes an intermediate step in the cumulant hierarchy. While for continuous random variables the higher-order cumulants are, by definition, independent of all lower ones, for discrete variables this is not necessarily the case: All cumulants of a random binary variable are completely determined by its mean. The constraint on only two possible values causes this dependence. The variance can therefore be determined by exploiting the binary character of the variables. A consequence of the lower-order moments depending only weakly on the statistics of higher-order moments is that cross-covariances can be determined from linear fluctuations around the steady state, which itself is determined by neglecting correlations altogether. We may therefore speculate that an approximation of third- or higher-order correlations can be obtained from the fluctuations around a state that itself is determined self-consistently by taking into account only up to second-order correlations.

The consistent truncation of the cumulant hierarchy further provides deeper insights: The expression for covariances between individual neuron pairs follows naturally from a consistent Gaussian closure, without further approximation. In previous studies, it was obtained as a linear approximation for weak correlations [see Eqs. (31)–(33) in Ref. [34]] and population-averaged covariances [see Eq. (6.8) of Ref. [33]]. Here, we show that solving the modified Lyapunov equation naturally leads to a decomposition of fluctuations into eigenmodes of the system. The

presented method solves the problem that the modified Lyapunov equation does not hold on the diagonal, in contrast to the case of population averages [33]. The approach moreover exposes that fluctuations in the Gaussian approximation are described by a set of coupled Ornstein-Uhlenbeck processes. It is well known that the fluctuations obtained from a systematic system-size expansion (see Chap. X of Ref. [72]), to lowest order, obey a Langevin equation. This result has been applied to networks of homogeneous populations [54,73,74]. The interesting point here is the feasibility of such a reduction directly on the level of individual binary variables. In contrast to the population-averaged activity, the jumps of a binary variable cannot be considered small compared to its value; hence, the precondition to apply the system-size expansion is not valid. The approach taken here to expose the correspondence between binary dynamics and Ornstein-Uhlenbeck processes is hence complementary. In particular, the result shows that the effective noises appearing in the equivalent set of Ornstein-Uhlenbeck processes are not uniquely determined by the activities of the neurons alone, in contrast to the population-averaged case (see Sec. 3.1 of Ref. [54]).

Decomposing externally applied signals to the network into the same eigenmodes as the intrinsically generated fluctuations, we obtain an expression for the susceptibility of the network on the single unit level that explains the spatiotemporal filtering applied to external signals, complementing the result for interacting populations (see Sec. V of Ref. [33]). The expression shows how the decay properties of the induced network response are determined by the eigenvalue associated with the eigendirections stimulated by the external signal. In particular, stimuli exciting modes close to an instability yield responses with a long memory lifetime. In Erdős-Rényi random binary networks with fixed weights, the spectral radius of the matrix of effective couplings is bounded by  $\sqrt{2(1-p)/\pi} < 0.8 < 1$  [59], so that all modes of the network are stable. An instability of the mean activities can only be achieved with coupling statistics or motifs that differ from an Erdős-Rényi network. This finding is in qualitative contrast to spiking networks that indeed show a rate instability at a critical coupling weight [75,76]. This insight has consequences for the use of random networks with fixed weights in the framework of reservoir computing [77,78], where highest computational performance [79] is achieved at the edge of chaos (for a review, see Ref. [80]).

We show that the effective interaction strength, i.e., the product of a unit's susceptibility and the incoming coupling amplitude, can uniquely be reconstructed from the covariance matrix and the slope of the covariance functions at zero time lag. This relation builds on the regression theorem [46], which states that the fluctuations in the system, to linear order (31), follow the same differential equation as the time-delayed covariance functions (2).

Coupling amplitudes, for principle reasons, cannot be inferred unambiguously: Correlations only depend on  $J/\sigma$ , where  $\sigma$  measures the strength of the overall incoming fluctuations to the node. The inference of the effective interaction strength  $J/\sigma$  still allows the classification of connections into excitatory and inhibitory ones, but it leads to a considerable rate of false-positive connections. The method presented here requires the measurement of covariances between all pairs of units in the network. In the case of severe subsampling, its application is restricted to small subnetworks. Thus, in general settings, more sophisticated approaches that take into account the influence of hidden units on the observed covariances [43], or on the inference of the global network state [81], are more promising. The presented expressions that relate the covariance matrix and its slope at zero time lag to the coupling matrix may still prove useful to construct similar inference methods for the investigated class of networks. Moreover, the algorithm constructing a network that generates activity with desired mean activities and covariances is a useful tool to generate surrogate data.

Previous works have addressed several aspects of pairwise correlations. The smallness of the average magnitude of covariances in strongly coupled balanced networks [44] has been explained by the influential work of Renart *et al.* [34] in the large- $N$  limit. In general, decorrelation follows from the dominant negative feedback in balanced networks at any network size [55,82]. These global properties are determined by collective fluctuations and can hence be described by population-averaged mean-field theory [33]. While such self-averaging observables in disordered systems can be obtained in ensembles of system realizations, here we discuss the statistics of individual units in a particular realization of the couplings, which goes beyond the existing approaches.

Our results are complementary to the method presented by Buice *et al.* [38], who consider Markov systems of interacting populations of neurons that consequently have integer numbers of active states, as opposed to the Glauber dynamics [3] of individual neurons considered here. In consequence, the truncation in the former work is performed on the level of normal-ordered cumulants, which represent an approximation around Poisson statistics, while here we derive an expansion in terms of ordinary cumulants [83]. Moreover, Buice *et al.* [38] and Ginzburg and Sompolinsky [33] consider all fluctuations beyond the mean activity perturbatively and therefore rely on a smooth activation function, while our results are also applicable to deterministic single units with hard thresholds.

In the path-integral formulation of Markovian neuronal network dynamics, introduced by Buice and Cowan [84], a systematic treatment of fluctuations without resorting to *ad hoc* approximations can be done by a loop-wise expansion (see Sec. 6.4 of Ref. [85]). The first correction term in this perturbative expansion corresponds to finite-size



fluctuations introduced by second-order cumulants, as discussed by Hildebrand, Buice, and Chow for coupled Kuramoto oscillators [86,87] and for more realistic spiking neuron models by Ref. [88]. The Gaussian approximation of  $h_k$  presented here is equivalent to a saddle-point approximation, the lowest order of the loop-wise expansion, in the auxiliary field [see, e.g., Eq. (3.5) of Ref. [28] or Eq. (3) of Ref. [29]]. The close-to-Gaussian approximation considers a subset of third-order cumulants of the activity variables as the first finite-size correction term. It neglects contributions to the third-order cumulant of the summed input  $h_i = \sum_j J_{ij} n_j$  that have three different unit indices. Only these cannot be expressed in terms of first and second cumulants of the involved units. This approximation scheme therefore includes all nontrivial dependencies between pairs of units ( $c_{ij}$ ) and excludes all nontrivial dependencies between three or more units consistently.

A different approach treating the statistics of individual units in a single network realization analytically originates in the spin-glass literature. Initiated by the symmetrically coupled Sherrington-Kirkpatrick spin-glass model [4], an expansion of the free energy in the coupling strength (Plefka expansion [89]) leads to the Thouless-Anderson-Palmer mean-field theory [5,12,90,91]. By construction, this method is restricted to weak coupling and, in its original form, starting from the partition function, also to systems in thermodynamic equilibrium. An extension to asymmetrically coupled nonequilibrium systems has been derived by invoking information-theoretic arguments [92]. A related approach has been taken by Ref. [41]. Still, the extended methods rely on the smoothness of the activation function and the smallness of the coupling constants, two restrictions that we are able to overcome here. Compared to previously mentioned works, the presented approach is more closely related to the mean-field theory by Mézard and Sakellariou [42]. Although they consider a system with a sequential Glauber update in discrete time steps, the method can be extended to the asynchronous update investigated here. Their analysis does not require weak coupling, similar to ours, but, in contrast, still relies on a smooth activation function for individual units. Moreover, their theory neglects the influence of the cross-covariance on the marginal statistics [their Eq. (4)], an important finite-size effect shown here to result in a distribution of mean activities due to correlations alone.

The mean-field methods employed in computer science, biology, artificial intelligence, social sciences, economics, and theoretical neuroscience (see, e.g., Refs. [31,32,93]) may be complemented by our results that go beyond population-averaged dynamics. The general formalism presented here, starting from the master equation, can be widely adapted by defining model-specific transition rates (see Table I of Ref. [36]) and particular types of couplings. The methods therefore also apply to equilibrium systems with Hamiltonian formulation, such as, e.g., associative

networks with parallel processing [94], and may be useful to selectively probe the energy landscape for ergodicity breaking, to characterize the statistical features of individual local energy minima, and to assess how many minima exist in these systems. In more general terms, the formalism enables the study of the dynamics of systems with arbitrary symmetry in the couplings, and it complements their analysis in terms of other features such as the storage capacity [95].

In neuroscience, with the advancement of electrophysiology, experimental data with hundreds of simultaneously recorded neurons have become readily available [96,97]. The availability of parallel data progressively changes the focus from the study of single cell responses to emergent phenomena arising through the interaction between neurons in networks [98]. Pairwise correlations are moreover closely linked to fluctuations of the population activity [82], which have been shown to shape experimentally accessible signals, such as the local field potential or the electroencephalogram (EEG) [99,100]. The effective description of the statistics of individual neurons, valid in the entire range of coupling strengths, forms the basis for studies of dynamics on adaptive networks [101], e.g., the interaction of neuronal dynamics with correlation-sensitive learning rules [102]. Explicit expressions, not only for zero time lag but also for the slope of covariance functions, allow the definition of plasticity rules resembling spike-timing dependent plasticity [25]. The finding that the collective dynamics is captured by the nontrivial second-order statistics implies that theoretical descriptions of mechanisms, e.g., biologically realistic synaptic learning, which usually only rely on first- and second-order statistics, have now come into reach. Balanced networks show widely distributed correlations across pairs of neurons with small mean [39]. This robust feature is captured by the presented linear equations for the covariance matrix. The shape of the distribution to date has not been related to the properties of the underlying network structure. Further work is required to obtain analytical expressions exposing how the structural properties give rise to the statistics of the distribution of covariances. Such results would enable us to deduce statistics of the connections from the observed activity. The presented description of the structure of fluctuations in this archetypical model of collective phenomena by a set of nonlinear equations provides a starting point in this endeavor and facilitates further development on disordered, coupled systems with large numbers of degrees of freedom.

## ACKNOWLEDGMENTS

This work was partially supported by the Helmholtz Young Investigator's Group VH-NG-1028, Helmholtz Portfolio Theme SMHB, and EU Grant No. 604102 (Human Brain Project, HBP). All simulations were carried out with NEST (<http://www.nest-initiative.org>).



## APPENDIX A: DERIVATION OF MOMENT EQUATIONS UP TO SECOND ORDER

For completeness and to establish a consistent notation, here we include the derivation of Eq. (2) for the first and second moments of the activity in a binary network. We follow the notation introduced in Buice *et al.* [38], which has been adapted to binary networks in Helias *et al.* [55].

The stochastic system is completely characterized by the joint probability distribution  $p(\mathbf{n})$  of all  $N$  binary variables  $\mathbf{n}$ . Knowing the joint probability distribution, arbitrary moments can be calculated, among them pairwise correlations. The occupation probability of each state follows the master equation (A1). We denote as  $\mathbf{n}_{i+} = (n_1, \dots, n_{i-1}, 1, n_{i+1}, \dots, n_N)$  the state, where the  $i$ th neuron is active ( $n_i = 1$ ), and  $\mathbf{n}_{i-}$  where neuron  $i$  is inactive ( $n_i = 0$ ). Since in each infinitesimal time interval at most one neuron can change state, for each given state  $\mathbf{n}$ , there are  $N$  possible transitions (each corresponding to one of the  $N$  neurons changing state). The sum of the probability fluxes into the state and out of the state must equal the change of probability in the respective state [9], i.e.,

$$\tau \frac{\partial p(\mathbf{n})}{\partial t} = \sum_{i=1}^N (2n_i - 1) (p(\mathbf{n}_{i-}) F_i(\mathbf{n}_{i-}) - p(\mathbf{n}_{i+}) (1 - F_i(\mathbf{n}_{i+}))) \quad \forall \mathbf{n} \in \{0, 1\}^N. \quad (\text{A1})$$

From this equation, we derive expressions for the first  $\langle n_k \rangle$  and second moments  $\langle n_k n_l \rangle$  by multiplying with  $n_k n_l$  and summing over all possible states  $\mathbf{n} \in \{0, 1\}^N$ , which leads to

$$\tau \frac{\partial}{\partial t} \langle n_k n_l \rangle = \sum_{\mathbf{n} \in \{0, 1\}^N} \sum_{i=1}^N n_k n_l (2n_i - 1) \times \underbrace{(p(\mathbf{n}_{i-}) F_i(\mathbf{n}_{i-}) - p(\mathbf{n}_{i+}) (1 - F_i(\mathbf{n}_{i+})))}_{\equiv G_i(\mathbf{n} \setminus n_i)},$$

where we denote as  $\langle f(\mathbf{n}) \rangle = \sum_{\mathbf{n} \in \{0, 1\}^N} p(\mathbf{n}) f(\mathbf{n})$  the average of a function  $f(\mathbf{n})$  with respect to the distribution  $p(\mathbf{n})$ . Note that the term denoted  $G_i(\mathbf{n} \setminus n_i)$  does not depend on the state of neuron  $i$ . We use the notation  $\mathbf{n} \setminus n_i$  for the state of the network excluding neuron  $i$ , i.e.,  $\mathbf{n} \setminus n_i = (n_1, \dots, n_{i-1}, n_{i+1}, \dots, n_N)$ . Separating the terms in the sum over  $i$  into those with  $i \neq k, l$  and the two terms with  $i = k$  and  $i = l$ , we obtain

$$\begin{aligned} \tau \frac{\partial}{\partial t} \langle n_k n_l \rangle &= \sum_{\mathbf{n}} \sum_{i=1, i \neq k, l}^N n_k n_l (2n_i - 1) G_i(\mathbf{n} \setminus n_i) \\ &\quad + n_k n_l (2n_k - 1) G_k(\mathbf{n} \setminus n_k) \\ &\quad + n_k n_l (2n_l - 1) G_l(\mathbf{n} \setminus n_l) \\ &= \sum_{i=1, i \neq k, l}^N \sum_{\mathbf{n} \setminus n_i} n_k n_l (G_i(\mathbf{n} \setminus n_i) - G_i(\mathbf{n} \setminus n_i)) \\ &\quad + \sum_{\mathbf{n}} n_k n_l G_k(\mathbf{n} \setminus n_k) + \sum_{\mathbf{n}} n_k n_l G_l(\mathbf{n} \setminus n_l), \end{aligned}$$

where we obtained the first term by explicitly summing over state  $n_i \in \{0, 1\}$  (i.e., using  $\sum_{\mathbf{n} \in \{0, 1\}^N} = \sum_{\mathbf{n} \setminus n_i \in \{0, 1\}^{N-1}} \sum_{n_i=0}^1$  and evaluating the sum  $\sum_{n_i=0}^1$ ). This first sum obviously vanishes. The remaining terms are of identical form, with the roles of  $k$  and  $l$  interchanged. We hence only consider the first of them and obtain the other by symmetry. The first term simplifies to

$$\begin{aligned} &\sum_{\mathbf{n}} n_k n_l G_k(\mathbf{n} \setminus n_k) \\ &\stackrel{n_k=1}{=} \sum_{\mathbf{n} \setminus n_k} n_l G_k(\mathbf{n} \setminus n_k) \\ &\stackrel{\text{def } G_k}{=} \begin{cases} \sum_{\mathbf{n} \setminus n_k} p(\mathbf{n}_{k-}) F_k(\mathbf{n}_{k-}) \\ \quad + p(\mathbf{n}_{k+}) F_k(\mathbf{n}_{k+}) - p(\mathbf{n}_{k+}) & \text{for } k = l \\ \sum_{\mathbf{n} \setminus n_k} p(\mathbf{n}_{k-}) n_l F_k(\mathbf{n}_{k-}) \\ \quad + p(\mathbf{n}_{k+}) n_l F_k(\mathbf{n}_{k+}) - n_l p(\mathbf{n}_{k+}) & \text{for } k \neq l \end{cases} \\ &= \begin{cases} \langle F_k(\mathbf{n}) \rangle - \langle n_k \rangle & \text{for } k = l \\ \langle F_k(\mathbf{n}) n_l \rangle - \langle n_k n_l \rangle & \text{for } k \neq l. \end{cases} \quad (\text{A2}) \end{aligned}$$

Taken together with the mirror term  $k \leftrightarrow l$ , we arrive at two conditions, one for the first ( $k = l$ ,  $\langle n_k^2 \rangle = \langle n_k \rangle$ ) and one for the second ( $k \neq l$ ) moment,

$$\tau \frac{\partial}{\partial t} \langle n_k \rangle \stackrel{k=l}{=} -\langle n_k \rangle + \langle F_k(\mathbf{n}) \rangle, \quad (\text{A3})$$

$$\tau \frac{\partial}{\partial t} \langle n_k n_l \rangle \stackrel{k \neq l}{=} -2\langle n_k n_l \rangle + \langle F_k(\mathbf{n}) n_l \rangle + \langle F_l(\mathbf{n}) n_k \rangle. \quad (\text{A4})$$

The time-lagged correlation function can be derived along completely analogous lines as Eq. (A4), as the forward time evolution equation (differential equation with respect to  $t$ ) of the two-point probability distribution  $p(\mathbf{n}, t, \mathbf{q}, s)$  fulfills, because of the Markov property, the same master equation (A1) as the equal-time probability distribution  $p(\mathbf{n}, t)$ . The resulting differential equation reads

$$\begin{aligned}\tau \frac{\partial}{\partial t} \langle n_k(t) n_l(s) \rangle &\equiv \tau \frac{\partial}{\partial t} \sum_{\mathbf{n}, \mathbf{q}} p(\mathbf{n}, t, \mathbf{q}, s) n_k q_l \\ &= -\langle n_k(t) n_l(s) \rangle + \langle F_k(\mathbf{n}(t)) n_l(s) \rangle.\end{aligned}\quad (\text{A5})$$

## APPENDIX B: CUMULANTS OF SUMMED RANDOM VARIABLES

Let

$$y = \sum_{i=1}^N x_i, \quad (\text{B1})$$

with  $x_i$  random variables that follow an arbitrary distribution  $p(\mathbf{x})$ . The moment generating functions of  $y$  and  $\mathbf{x}$  are then related by

$$\begin{aligned}\varphi_y(t) &= \sum_{\mathbf{x}} p(\mathbf{x}) e^{t y} = \langle e^{t y} \rangle_{\mathbf{x}} \\ &= \langle e^{t \sum_{i=1}^N x_i} \rangle_{\mathbf{x}} \\ &= (\varphi_{\mathbf{x}} \circ \iota)(t),\end{aligned}\quad (\text{B2})$$

where concatenation with the function  $\iota: t \mapsto (t, \dots, t)$  replaces every  $t_i$  by a  $t$ . The cumulant generating functions therefore follow as

$$\begin{aligned}\Phi_y(t) &= \ln \varphi_y(t) \\ &= \ln (\varphi_{\mathbf{x}} \circ \iota)(t) = (\Phi_{\mathbf{x}} \circ \iota)(t),\end{aligned}\quad (\text{B3})$$

with  $\Phi_{\mathbf{x}}(\mathbf{t}) = \ln \langle e^{\mathbf{t} \cdot \mathbf{x}} \rangle$  the cumulant generating function of  $\mathbf{x}$  and  $\mathbf{t} \cdot \mathbf{x}$  the scalar product. From the expansion in cumulants  $\kappa_{ij}^{\mathbf{x}}$  for  $\mathbf{x}$  and  $\kappa_{1,2,\dots}$  for  $y$ , we get the following relationship:

$$\begin{aligned}\Phi_y(t) &\equiv \sum_{l=1}^{\infty} \frac{\kappa_l}{l!} t^l \\ &= (\Phi_{\mathbf{x}} \circ \iota)(t) \\ &= \left[ \left( \sum_{i=1}^N \kappa_i^{\mathbf{x}} t_i + \sum_{i,j=1}^N \frac{\kappa_{ij}^{\mathbf{x}}}{2!} t_i t_j + \dots \right) \circ \iota \right](t) \\ &= \underbrace{\sum_{i=1}^N \kappa_i^{\mathbf{x}} t}_{\kappa_1} + \frac{1}{2!} \underbrace{\sum_{i,j=1}^N \kappa_{ij}^{\mathbf{x}} t^2}_{\kappa_2} + \dots,\end{aligned}\quad (\text{B4})$$

so the cumulants of the summed variable  $\kappa_1, \kappa_2$ , etc. are given by the sums of the cumulants of the individual variables of corresponding order.

## APPENDIX C: FUNCTIONS OF CLOSE-TO-GAUSSIAN VARIABLES

To determine the mean activity, we need to apply a nonlinear function  $f$  to a variable  $h_k$  that has a statistics close to Gaussian. For brevity, we suppress the index  $k$  in

the following. We assume that the Fourier transform  $\hat{f}(\omega)$  exists. Let  $y$  be the random variable (B1), which is the sum of a large number of individual variables  $x_i$  and is assumed to be close to Gaussian. For the expectation value  $\langle f(y) \rangle_{\mathbf{x}}$ , we get

$$\begin{aligned}\langle f(y) \rangle_{\mathbf{x}} &= \frac{1}{2\pi} \int d\omega \hat{f}(\omega) \langle e^{i\omega y} \rangle_{\mathbf{x}} \\ &= \frac{1}{2\pi} \int d\omega \hat{f}(\omega) e^{\Phi_y(i\omega)} \\ &= \frac{1}{2\pi} \int d\omega \hat{f}(\omega) e^{\kappa_1(i\omega) + \frac{1}{2!}\kappa_2(i\omega)^2 + \frac{1}{3!}\kappa_3(i\omega)^3 + \dots},\end{aligned}\quad (\text{C1})$$

where  $\Phi_y$  is the cumulant generating function of  $y$  [Eq. (B3)] and  $\kappa_i$  denotes the  $i$ th cumulant of  $y$ . We are interested in an approximation that treats the dominant first and second (Gaussian) cumulants of  $y$  explicitly and separate the effect of all higher cumulants by writing

$$\begin{aligned}\langle f(y) \rangle_{\mathbf{x}} &= e^{\frac{1}{6}\kappa_3(\partial/\partial\kappa_1)^3 + \dots} \int d\omega \hat{f}(\omega) e^{\kappa_1(i\omega) + \frac{1}{2}\kappa_2(i\omega)^2} \\ &= e^{\frac{1}{6}\kappa_3(\partial/\partial\kappa_1)^3 + \dots} \langle f(y) \rangle_{y \sim \mathcal{N}(\kappa_1, \kappa_2)},\end{aligned}\quad (\text{C2})$$

where we identified, in the last line, the Fourier transform of a Gaussian with moments  $\kappa_1$  and  $\kappa_2$  via Eq. (C1).

For the covariance, we need to evaluate terms of the form  $\langle F_k n_l \rangle$ . These terms can be obtained analogously as

$$\begin{aligned}\langle f(y) x_l \rangle_{\mathbf{x}} &= \frac{1}{2\pi} \int d\omega \hat{f}(\omega) \langle x_l e^{i\omega y} \rangle_{\mathbf{x}} \\ &= \frac{1}{2\pi} \int d\omega \hat{f}(\omega) ((\partial_{t_l} \varphi_{\mathbf{x}}) \circ \iota)(i\omega) \\ &= \frac{1}{2\pi} \int d\omega \hat{f}(\omega) (([\partial_{t_l} \Phi_{\mathbf{x}}] e^{\Phi_{\mathbf{x}}}) \circ \iota)(i\omega),\end{aligned}$$

where  $\varphi_{\mathbf{x}}$  is the moment generating function of  $\mathbf{x}$  [Eq. (B2)]. The derivative of the cumulant generating function with Eq. (B4) becomes

$$\begin{aligned}\partial_{t_l} \Phi_{\mathbf{x}}(\mathbf{t}) &= \kappa_l^{\mathbf{x}} + \sum_{j=1}^N \kappa_{lj}^{\mathbf{x}} t_j + \frac{1}{2} \sum_{i,j=1}^N \kappa_{lij}^{\mathbf{x}} t_i t_j + \dots, \\ (\partial_{t_l} \Phi_{\mathbf{x}} \circ \iota)(i\omega) &= \underbrace{\kappa_l^{\mathbf{x}}}_{=\Delta\kappa_{1,l}} + \sum_{j=1}^N \underbrace{\kappa_{lj}^{\mathbf{x}}(i\omega)}_{=\Delta\kappa_{2,l}} + \frac{1}{2} \sum_{i,j=1}^N \underbrace{\kappa_{lij}^{\mathbf{x}}(i\omega)^2}_{=\Delta\kappa_{3,l}} + \dots \\ &= \sum_{q=0}^{\infty} \Delta\kappa_{q,l}(i\omega)^q,\end{aligned}$$

where the product rule produced a factor 2 in the second term and a factor 3 in the third term, and we used the symmetry of the cumulants with respect to permutations of indices. So we get

$$\begin{aligned}
\langle x_l f(y) \rangle_{\mathbf{x}} &= \int d\omega \hat{f}(\omega) \sum_{q=0}^{\infty} \Delta \kappa_{q,l} (i\omega)^q e^{\sum_{p=0}^{\infty} \kappa_p (i\omega)^p} \\
&= \left[ \sum_{q=0}^{\infty} \Delta \kappa_{q,l} \frac{\partial}{\partial \kappa_q} \right] e^{(1/3!) \kappa_3 (\partial/\partial \kappa_1)^3 + \dots} \langle f(y) \rangle_{y \sim \mathcal{N}(\kappa_1, \kappa_2)} \\
&= \left[ \sum_{q=0}^{\infty} \Delta \kappa_{q,l} \frac{1}{q!} \left( \frac{\partial}{\partial \kappa_1} \right)^q \right] \langle f(y) \rangle_{\mathbf{x}},
\end{aligned} \tag{C3}$$

where we replaced  $(i\omega)^q$  by derivatives of the exponential with respect to cumulants and identified  $\langle f(y) \rangle_{\mathbf{x}}$  in the last step using Eq. (C2).

#### APPENDIX D: TRIVIAL THIRD- AND FOURTH-ORDER CUMULANTS OF BINARY VARIABLES EXPRESSED BY LOWER-ORDER CUMULANTS

*Third order.*—Let  $n_l, n_i, n_j, n_r \in [0, 1]$  be binary variables. The raw third moment can be written as a sum of all combinations of cumulants up to order three [Eq. (20)],  $\langle n_l n_i n_j \rangle = \langle\langle n_l n_i n_j \rangle\rangle + c_{li} m_j + c_{ij} m_l + c_{jl} m_i + m_l m_i m_j$ . Using  $n_i^K = n_i$  for each integer  $K \geq 1$ , in the case of binary variables  $n_i$ , we consider the two cases

$$\begin{aligned}
l = i \neq j: \quad \langle n_l n_j \rangle &= \langle\langle n_l n_i n_j \rangle\rangle + c_{li} m_j + c_{ij} m_l + c_{jl} m_i + m_l^2 m_j, \\
\langle\langle n_l n_i n_j \rangle\rangle &= \underbrace{\langle n_l n_j \rangle}_{c_{lj} + \langle n_l \rangle \langle n_j \rangle} - c_{li} m_j - 2c_{lj} m_l - m_l^2 m_j \\
&= c_{lj} (1 - 2m_l) + \underbrace{(-c_{li} + m_l - m_l^2)}_{=0} m_j \\
&= c_{lj} (1 - 2m_l), \\
l = i = j: \quad \langle n_l \rangle &= \langle\langle n_l n_i n_l \rangle\rangle + 3c_{li} m_l + m_l^3, \\
\langle\langle n_l n_i n_l \rangle\rangle &= m_l - 3m_l (1 - m_l) m_l - m_l^3 \\
&= m_l - 3m_l^2 + 2m_l^3 \\
&= c_{li} (1 - 2m_l),
\end{aligned} \tag{D1}$$

which together yield the expression (21) in the main text.

The third cumulant [Eq. (17)] of  $h_k$  follows from the assumption  $\langle\langle n_i n_j n_r \rangle\rangle \simeq 0$  for  $i \neq j \neq r$  as

$$\begin{aligned}
\kappa_{3,k} &= \sum_{ijr} J_{ki} J_{kj} J_{kr} \langle\langle n_i n_j n_r \rangle\rangle \\
&= \underbrace{\sum_{i \neq j \neq r} J_{ki} J_{kj} J_{kr} \langle\langle n_i n_j n_r \rangle\rangle}_{N(N-1)(N-2) \text{ terms}} + 3 \underbrace{\sum_{i=j \neq r} J_{ki}^2 J_{kr} \langle\langle n_i n_i n_r \rangle\rangle}_{N(N-1) \text{ terms}} + \underbrace{\sum_{i=j=r} J_{ki}^3 \langle\langle n_i n_i n_i \rangle\rangle}_N \\
&\simeq 3 \sum_{i=j \neq r} J_{ki}^2 J_{kr} c_{ir} (1 - 2m_i) + \sum_{i=j=r} J_{ki}^3 c_{ii} (1 - 2m_i) \\
&= 3 \sum_{i,r} J_{ki}^2 (1 - 2m_i) c_{ir} J_{kr} - 2 \sum_i J_{ki}^3 c_{ii} (1 - 2m_i),
\end{aligned} \tag{D2}$$

where we included the term  $i = r$  in the first sum and compensated accordingly in the latter term. With  $c_{ii} (1 - 2m_i) = m_i - 3m_i^2 + 2m_i^3$ , we obtain the expression (22) in the main text. Analogously, it follows that

$$\begin{aligned}
\sum_{i,j=1}^N J_{ki}J_{kj}\langle\langle n_in_in_j\rangle\rangle &= \underbrace{\sum_{i\neq j\neq l} J_{ki}J_{kj}\langle\langle n_in_in_j\rangle\rangle}_{\approx 0} + \sum_{l\neq i=j=1}^N J_{ki}^2\langle\langle n_in_in_i\rangle\rangle + \sum_{i=l\neq j=1}^N J_{kl}J_{kj}\langle\langle n_in_in_j\rangle\rangle \\
&+ \sum_{j=l\neq i=1}^N J_{ki}J_{kl}\langle\langle n_in_in_l\rangle\rangle + J_{kl}^2\langle\langle n_in_in_l\rangle\rangle \\
&= \sum_{l\neq i=1}^N J_{ki}^2c_{li}(1-2m_i) + 2J_{kl}\sum_{l\neq i=1}^N J_{ki}c_{li}(1-2m_l) + J_{kl}^2c_{ll}(1-2m_l) \\
&= \sum_{i=1}^N J_{ki}^2(1-2m_i)c_{li} + 2J_{kl}(1-2m_l)\sum_{i=1}^N J_{ki}c_{li} - 2J_{kl}^2c_{ll}(1-2m_l), \tag{D3}
\end{aligned}$$

which yields the expression (23) in the main text.

*Fourth order.*—The cumulant of fourth order is

$$\begin{aligned}
\langle n_in_jn_rn_l\rangle &= \langle\langle n_in_jn_rn_l\rangle\rangle + \langle\langle n_in_jn_r\rangle\rangle m_l + \langle\langle n_jn_rn_l\rangle\rangle m_i + \langle\langle n_rn_in_i\rangle\rangle m_j + \langle\langle n_in_in_j\rangle\rangle m_r + c_{ij}m_rm_l + c_{ir}m_jm_l + c_{il}m_rm_j \\
&+ c_{jr}m_lm_i + c_{jl}m_rm_i + c_{rl}m_im_j + c_{ij}c_{rl} + c_{ir}c_{jl} + c_{il}c_{rj} + m_im_jm_rm_l.
\end{aligned}$$

Only those cumulants in which at most two different indices appear are fixed by first- and second-order cumulants. We need to distinguish three cases. The first case is  $i = j \neq r = l$  and leads, with Eq. (D1), to the matrix

$$\begin{aligned}
\{\langle\langle n_in_in_rn_r\rangle\rangle_{ir}\} &= \mathbf{C} \otimes (\mathbf{1} - 2\mathbf{C}) - \text{diag}(\mathbf{C})\text{diag}(\mathbf{C})^T - 2\text{diag}(\mathbf{1} - 2\mathbf{m})\mathbf{C}\text{diag}(\mathbf{m}) - 2\text{diag}(\mathbf{m})\mathbf{C}\text{diag}(\mathbf{1} - 2\mathbf{m}) \\
&- \text{diag}(\mathbf{C})(\mathbf{m} \otimes \mathbf{m})^T - (\mathbf{m} \otimes \mathbf{m})\text{diag}(\mathbf{C})^T - 4\text{diag}(\mathbf{m})\mathbf{C}\text{diag}(\mathbf{m}) + \mathbf{m}\mathbf{m}^T \otimes (\mathbf{1} - \mathbf{m}\mathbf{m}^T). \tag{D4}
\end{aligned}$$

The second case  $i = j = l \neq r$  yields the matrix

$$\begin{aligned}
\{\langle\langle n_in_in_rn_r\rangle\rangle_{ir}\} &= \text{diag}(\mathbf{1} - 3\text{diag}(\mathbf{C}))\mathbf{C} - \text{diag}(\mathbf{C})\mathbf{m}^T - 3\text{diag}(\mathbf{m})\mathbf{C} - (\mathbf{m} \otimes \text{diag}(\mathbf{C}))\mathbf{m}^T + 3\text{diag}(\mathbf{m} \otimes \mathbf{m})\mathbf{C} \\
&+ \text{diag}(\mathbf{1} - \mathbf{m} \otimes \mathbf{m})\mathbf{m}\mathbf{m}^T, \tag{D5}
\end{aligned}$$

and the third for  $i = j = r = l$  yields a vector

$$\langle\langle n_in_in_in_i\rangle\rangle = \mathbf{m} \otimes (\mathbf{1} - (\mathbf{m} \otimes)^3) - 4\text{diag}(\mathbf{C}) \otimes \mathbf{m} + 2\text{diag}(\mathbf{C}) \otimes (\mathbf{m} \otimes)^2 - 3(\text{diag}(\mathbf{C}) \otimes)^2, \tag{D6}$$

where we use the notation  $(\mathbf{x} \otimes)^n = \mathbf{x} \otimes \dots \otimes \mathbf{x}$  for the element-wise  $n$ th power of a vector. In Eq. (19), we need the term

$$\begin{aligned}
\Delta\kappa_{3,kl} &= \sum_{i,j,r=1}^N J_{ki}J_{kj}J_{kr}\langle\langle n_in_jn_rn_l\rangle\rangle \\
&\approx 3\sum_i J_{ki}^2J_{kl}\langle\langle n_in_in_in_l\rangle\rangle \quad (i = j, r = l), (i = r, j = l), (i = l, j = r) \\
&+ \sum_i J_{ki}^3\langle\langle n_in_in_in_i\rangle\rangle \quad (i = j = r \neq l) \\
&+ 3\sum_i J_{ki}J_{kl}^2\langle\langle n_in_in_in_i\rangle\rangle \quad (j = r = l \neq i), (r = l = i \neq j), (l = i = j \neq r) \\
&+ J_{kl}^3\langle\langle n_in_in_in_l\rangle\rangle \quad (i = j = r = l),
\end{aligned}$$

which in matrix form gives rise to expression (24) of the main text.



- [1] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. Newman, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group* (Oxford University Press, New York, 1992).
- [2] E. Ising, *Beitrag zur Theorie des Ferromagnetismus*, *Z. Phys.* **31**, 253 (1925).
- [3] R. Glauber, *Time-Dependent Statistics of the Ising Model*, *J. Math. Phys. (N.Y.)* **4**, 294 (1963).
- [4] D. Sherrington and S. Kirkpatrick, *Solvable Model of a Spin-Glass*, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [5] D. J. Thouless, P. W. Anderson, and R. G. Palmer, *Solution of "Solvable Model of a Spin Glass"*, *Philos. Mag.* **35**, 593 (1977).
- [6] J. Hertz, G. Grinstein, and S. Solla, *Memory Networks with Asymmetric Bonds*, *AIP Conf. Proc.* **151**, 212 (1986).
- [7] B. Derrida, *Dynamical Phase Transition in Non-symmetric Spin Glasses*, *J. Phys. A* **20**, L721 (1987).
- [8] A. Crisanti and H. Sompolinsky, *Dynamics of Spin Systems with Randomly Asymmetric Bonds: Langevin Dynamics and a Spherical Model*, *Phys. Rev. A* **36**, 4922 (1987).
- [9] F. Kelly, *Reversibility and Stochastic Networks* (Cambridge University Press, Cambridge, England, 1979).
- [10] U. Seifert, *Stochastic Thermodynamics, Fluctuation Theorems, and Molecular Machines*, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [11] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond*, *World Scientific Lecture Notes in Physics* (World Scientific, Singapore, 1987), Vol. 9.
- [12] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Clarendon Press, Oxford, 2001).
- [13] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).
- [14] C. Castellano, S. Fortunato, and V. Loreto, *Statistical Physics of Social Dynamics*, *Rev. Mod. Phys.* **81**, 591 (2009).
- [15] D. L. Stein and C. M. Newman, *Spin Glasses and Complexity* (Princeton University Press, Princeton, 2013).
- [16] D. Sornette, *Physics and Financial Economics (1776–2014): Puzzles, Ising and Agent-Based Models*, *Rep. Prog. Phys.* **77**, 062001 (2014).
- [17] J. C. Eccles, P. Fatt, and K. Koketsu, *Cholinergic and Inhibitory Synapses in a Pathway from Motor-Axon Collaterals to Motoneurons*, *J. Physiol.* **126**, 524 (1954).
- [18] Z. Li and P. Dayan, *Computational Differences between Asymmetrical and Symmetrical Networks*, *Network: Computation in Neural Systems* **10**, 59 (1999).
- [19] M. R. Cohen and A. Kohn, *Measuring and Interpreting Neuronal Correlations*, *Nat. Rev. Neurosci.* **14**, 811 (2011).
- [20] E. Zohary, M. N. Shadlen, and W. T. Newsome, *Correlated Neuronal Discharge Rate and Its Implications for Psychophysical Performance*, *Nature (London)* **370**, 140 (1994).
- [21] M. Shamir and H. Sompolinsky, in *Advances in Neural Information Processing Systems* (MIT Press, Cambridge, MA, 2001), pp. 277–284.
- [22] B. E. Kilavik, S. Roux, A. Ponce-Alvarez, J. Confais, S. Gruen, and A. Riehle, *Long-Term Modifications in Motor Cortical Dynamics Induced by Intensive Practice*, *J. Neurosci.* **29**, 12653 (2009).
- [23] H. Markram, J. Lübke, M. Frotscher, and B. Sakmann, *Regulation of Synaptic Efficacy by Coincidence of Post-synaptic APs and EPSPs*, *Science* **275**, 213 (1997).
- [24] G. Bi and M. Poo, *Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type*, *J. Neurosci.* **18**, 10464 (1998).
- [25] A. Morrison, M. Diesmann, and W. Gerstner, *Phenomenological Models of Synaptic Plasticity Based on Spike-Timing*, *Biol. Cybern.* **98**, 459 (2008).
- [26] J. Fiser, P. Berkes, G. Orbán, and M. Lengyel, *Statistically Optimal Perception and Learning: From Behavior to Neural Representations*, *Theor. Comput. Sci.* **14**, 119 (2010).
- [27] H. Sompolinsky and A. Zippelius, *Dynamic Theory of the Spin-Glass Phase*, *Phys. Rev. Lett.* **47**, 359 (1981).
- [28] H. Sompolinsky and A. Zippelius, *Relaxational Dynamics of the Edwards-Anderson Model and the Mean-Field Theory of Spin-Glasses*, *Phys. Rev. B* **25**, 6860 (1982).
- [29] H. Sompolinsky, A. Crisanti, and H. J. Sommers, *Chaos in Random Neural Networks*, *Phys. Rev. Lett.* **61**, 259 (1988).
- [30] L. F. Cugliandolo and J. Kurchan, *Analytical Solution of the Off-Equilibrium Dynamics of a Long-Range Spin-Glass Model*, *Phys. Rev. Lett.* **71**, 173 (1993).
- [31] J. Aljadeff, M. Stern, and T. Sharpee, *Transition to Chaos in Random Networks with Cell-Type-Specific Connectivity*, *Phys. Rev. Lett.* **114**, 088101 (2015).
- [32] J. Kadmon and H. Sompolinsky, *Transition to Chaos in Random Neuronal Networks*, *Phys. Rev. X* **5**, 041030 (2015).
- [33] I. Ginzburg and H. Sompolinsky, *Theory of Correlations in Stochastic Neural Networks*, *Phys. Rev. E* **50**, 3171 (1994).
- [34] A. Renart, J. De La Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris, *The Asynchronous State in Cortical Circuits*, *Science* **327**, 587 (2010).
- [35] J. P. Gleeson, *High-Accuracy Approximation of Binary-State Dynamics on Networks*, *Phys. Rev. Lett.* **107**, 068701 (2011).
- [36] J. P. Gleeson, *Binary-State Dynamics on Complex Networks: Pair Approximation and Beyond*, *Phys. Rev. X* **3**, 021004 (2013).
- [37] A. C. C. Coolen, S. N. Laughton, and D. Sherrington, *Dynamical Replica Theory for Disordered Spin Systems*, *Phys. Rev. B* **53**, 8184 (1996).
- [38] M. A. Buice, J. D. Cowan, and C. C. Chow, *Systematic Fluctuation Expansion for Neural Network Activity Equations*, *Neural Comput.* **22**, 377 (2010).
- [39] A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, and N. K. Logothetis, *Decorrelated Neuronal Firing in Cortical Microcircuits*, *Science* **327**, 584 (2010).
- [40] A. Aertsens and H. Preißl, in *Nonlinear Dynamics and Neuronal Networks*, *Proceedings of the 63rd W. E. Heraeus Seminar Friedrichsdorf 1990*, edited by H. G. Schuster (Wiley-VCH, Weinheim, 1990), pp. 281–301.
- [41] H.-L. Zeng, E. Aurell, M. Alava, and H. Mahmoudi, *Network Inference Using Asynchronously Updated Kinetic Ising Model*, *Phys. Rev. E* **83**, 041135 (2011).

- [42] M. Mézard and J. Sakellariou, *Exact Mean-Field Inference in Asymmetric Kinetic Ising Systems*, *J. Stat. Mech.* (2011) L07001.
- [43] J. Tyrcha and J. Hertz, *Network inference with hidden units*, [arXiv:1301.7274v1](#).
- [44] C. van Vreeswijk and H. Sompolinsky, *Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity*, *Science* **274**, 1724 (1996).
- [45] G. E. Uhlenbeck and L. S. Ornstein, *On the Theory of the Brownian Motion*, *Phys. Rev.* **36**, 823 (1930).
- [46] H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, Heidelberg, 1996).
- [47] D. Willshaw, O. Buneman, and H. Longuet-Higgins, *Non Holographic Associative Memory*, *Nature (London)* **222**, 960 (1969).
- [48] S.-I. Amari, *Homogeneous Nets of Neuron-like Elements*, *Biol. Cybern.* **17**, 211 (1975).
- [49] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Perseus Books Group, New York, 1991).
- [50] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing, Explorations in the Microstructure of Cognition: Foundations* (MIT Press, Cambridge, MA, 1986), Vol. 1.
- [51] J. J. Hopfield, *Neural Networks and Physical Systems with Emergent Collective Computational Abilities*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).
- [52] A. Hanuschkin, S. Kunkel, M. Helias, A. Morrison, and M. Diesmann, *A General and Efficient Method for Incorporating Precise Spike Times in Globally Time-Driven Simulations*, *Front. Neuroinform.* **4**, 113 (2010).
- [53] J. M. Eppler, R. Pauli, A. Peyser, T. Ippen, A. Morrison, J. Senk, W. Schenck, H. Bos, M. Helias, M. Schmidt *et al.*, *NEST 2.8.0* (2015).
- [54] D. Grytskyy, T. Tetzlaff, M. Diesmann, and M. Helias, *A Unified View on Weakly Correlated Recurrent Networks*, *Front. Comput. Neurosci.* **7**, 131 (2013).
- [55] M. Helias, T. Tetzlaff, and M. Diesmann, *The Correlation Structure of Local Cortical Networks Intrinsically Results from Recurrent Dynamics*, *PLoS Comput. Biol.* **10**, e1003428 (2014).
- [56] C. van Vreeswijk and H. Sompolinsky, *Chaotic Balanced State in a Model of Cortical Circuits*, *Neural Comput.* **10**, 1321 (1998).
- [57] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd ed., Springer Series in Synergetics No. 13 (Springer-Verlag, Berlin, 1985).
- [58] C. Van Vreeswijk and H. Sompolinsky, *Chaotic Balanced State in a Model of Cortical Circuits*, *Neural Comput.* **10**, 1321 (1998).
- [59] D. Grytskyy, T. Tetzlaff, M. Diesmann, and M. Helias, *Invariance of Covariances Arises out of Noise*, *AIP Conf. Proc.* **1510**, 258 (2013).
- [60] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Elsevier Science Publishers, Amsterdam, 1992).
- [61] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation* (John Wiley & Sons, New York, 2009).
- [62] *NIST Handbook of Mathematical Functions*, edited by F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (Cambridge University Press, New York, NY, 2010).
- [63] S. J. van Albada, M. Helias, and M. Diesmann, *Scalability of Asynchronous Networks Is Limited by One-to-One Mapping between Effective Connectivity and Correlations*, *PLoS Comput. Biol.* **11**, e1004490 (2015).
- [64] S. Cocco, R. Monasson, and V. Sessak, *High-Dimensional Inference with the Generalized Hopfield Model: Principal Component Analysis and Corrections*, *Phys. Rev. E* **83**, 051123 (2011).
- [65] Y. Roudi, E. Aurell, and J. A. Hertz, *Statistical Physics of Pairwise Probability Models*, *Front. Comput. Neurosci.* **3**, 22 (2009).
- [66] T. Mora, A. M. Walczak, W. Bialek, and C. G. Callan, *Maximum Entropy Models for Antibody Diversity*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5405 (2010).
- [67] G. Hermann and J. Touboul, *Heterogeneous Connections Induce Oscillations in Large-Scale Networks*, *Phys. Rev. Lett.* **109**, 018702 (2012).
- [68] S. Shinomoto, *A Cognitive and Associative Memory*, *Biol. Cybern.* **57**, 197 (1987).
- [69] J. P. Roach, L. M. Sander, and M. R. Zochowski, *Memory Recall and Spike Frequency Adaptation*, *Phys. Rev. E* **93**, 052307 (2016).
- [70] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, *Weak Pairwise Correlations Imply Strongly Correlated Network States in a Neural Population*, *Nature (London)* **440**, 1007 (2006).
- [71] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Perseus Books Group, New York, 1998).
- [72] N. van Kampen, *Stochastic Processes in Chemistry and Physics* (North-Holland, Amsterdam, 1992).
- [73] T. Ohira and J. Cowan, *Mathematics of Neural Networks*, edited by S. W. Ellacott, J. C. Mason, and I. J. Anderson (Springer, New York, 1997), pp. 290–294.
- [74] P. C. Bressloff, *Stochastic Neural Field Theory and the System-Size Expansion*, *SIAM J. Appl. Math.* **70**, 1488 (2009).
- [75] S. Ostojic, *Two Types of Asynchronous Activity in Networks of Excitatory and Inhibitory Spiking Neurons*, *Nat. Neurosci.* **17**, 594 (2014).
- [76] O. Harish and D. Hansel, *Asynchronous Rate Chaos in Spiking Neuronal Circuits*, *PLoS Comput. Biol.* **11**, e1004266 (2015).
- [77] H. Jaeger, *The “Echo State” Approach to Analysing and Training Recurrent Neural Networks*, GMD Report 148 (German National Research Center for Information Technology, St. Augustin, Germany, 2001).
- [78] W. Maass, T. Natschläger, and H. Markram, *Real-Time Computing without Stable States: A New Framework for Neural Computation Based on Perturbation*, *Neural Comput.* **14**, 2531 (2002).
- [79] J. P. Crutchfield and K. Young, *Inferring Statistical Complexity*, *Phys. Rev. Lett.* **63**, 105 (1989).
- [80] R. Legenstein and W. Maass, *What Makes a Dynamical System Computationally Powerful?* in *New Directions in*

- Statistical Signal Processing: From System to Brains*, edited by S. Haykin, J. C. Principe, T. J. Sejnowski, and J. G. McWhirter (MIT Press, Cambridge, MA, 2007), pp. 127–154.
- [81] M. A. Buice and C. C. Chow, *Effective Stochastic Behavior in Dynamical Systems with Incomplete Information*, *Phys. Rev. E* **84**, 051120 (2011).
  - [82] T. Tetzlaff, M. Helias, G. Einevoll, and M. Diesmann, *Decorrelation of Neural-Network Activity by Inhibitory Feedback*, *PLoS Comput. Biol.* **8**, e1002596 (2012).
  - [83] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd ed., Springer Series in Synergetics (Springer, New York, 2004).
  - [84] M. A. Buice and J. D. Cowan, *Field-Theoretic Approach to Fluctuation Effects in Neural Networks*, *Phys. Rev. E* **75**, 051919 (2007).
  - [85] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon Press, Oxford, 1996).
  - [86] E. J. Hildebrand, M. A. Buice, and C. C. Chow, *Kinetic Theory of Coupled Oscillators*, *Phys. Rev. Lett.* **98**, 054101 (2007).
  - [87] M. A. Buice and C. C. Chow, *Correlations, Fluctuations, and Stability of a Finite-Size Network of Coupled Oscillators*, *Phys. Rev. E* **76**, 031118 (2007).
  - [88] M. A. Buice and C. C. Chow, *Dynamic Finite Size Effects in Spiking Neural Networks*, *PLoS Comput. Biol.* **9**, e1002872 (2013).
  - [89] T. Plefka, *Convergence Condition of the Tap Equation for the Infinite-Ranged Ising Spin Glass Model*, *J. Phys. A* **15**, 1971 (1982).
  - [90] K. Nakanishi and H. Takayama, *Mean-Field Theory for a Spin-Glass Model of Neural Networks: Tap Free Energy and the Paramagnetic to Spin-Glass Transition*, *J. Phys. A* **30**, 8085 (1997).
  - [91] T. Tanaka, *Mean-Field Theory of Boltzmann Machine Learning*, *Phys. Rev. E* **58**, 2302 (1998).
  - [92] H. J. Kappen and J. J. Spanjers, *Mean Field Theory for Asymmetric Neural Networks*, *Phys. Rev. E* **61**, 5658 (2000).
  - [93] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, *Is the Voter Model a Model for Voters?*, *Phys. Rev. Lett.* **112**, 158701 (2014).
  - [94] P. Sollich, D. Tantari, A. Annibale, and A. Barra, *Extensive Parallel Processing on Scale-Free Networks*, *Phys. Rev. Lett.* **113**, 238106 (2014).
  - [95] A. Alemi, C. Baldassi, N. Brunel, and R. Zecchina, *A Three-Threshold Learning Rule Approaches the Maximal Capacity of Recurrent Neural Networks*, *PLoS Comput. Biol.* **11**, e1004439 (2015).
  - [96] G. Buzsáki, *Large-Scale Recording of Neuronal Ensembles*, *Nat. Neurosci.* **7**, 446 (2004).
  - [97] E. N. Brown, R. E. Kaas, and P. P. Mitra, *Multiple Neural Spike Train Data Analysis: State-of-the-Art and Future Challenges*, *Nat. Neurosci.* **7**, 456 (2004).
  - [98] I. H. Stevenson and K. P. Kording, *How Advances in Neural Recording Affect Data Analysis*, *Nat. Neurosci.* **14**, 139 (2011).
  - [99] H. Lindén, T. Tetzlaff, T. C. Potjans, K. H. Pettersen, S. Grün, M. Diesmann, and G. T. Einevoll, *Modeling the Spatial Reach of the LFP*, *Neuron* **72**, 859 (2011).
  - [100] A. Mazzoni, K. Whittingstall, N. Brunel, N. K. Logothetis, and S. Panzeri, *Understanding the Relationships between Spike Rate and Delta/Gamma Frequency Bands of LFPs and EEGs Using a Local Cortical Network Model*, *NeuroImage* **52**, 956 (2010).
  - [101] T. Gross, C. J. D. D’Lima, and B. Blasius, *Epidemic Dynamics on an Adaptive Network*, *Phys. Rev. Lett.* **96**, 208701 (2006).
  - [102] G. K. Ocker, A. Litwin-Kumar, and B. Doiron, *Self-Organization of Microcircuits in Networks of Spiking Neurons with Plastic Synapses*, *PLoS Comput. Biol.* **11**, e1004458 (2015).